

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

101 GCCCTGGCAGCGGTAGCCAGATTGAGCTAGCGCTGCGTCCGGGT 147
|||||
34 rgrprogllyserglyserclnlsprtsrtyrglyleuargprogly 49

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-694-865-4

seq_documentation_block:

; Sequence 4, Application US/08694865
; Patent No. 5837268
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,865
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-694-865-4

alignment_scores:
Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-306-689-10 x US-08-694-865-4 ..

Align seg 1/1 to: US-08-694-865-4 from: 1 to: 49

1 CAGATTGAGCTAGGCGCTGCGCTGGCAGCGGTTCTAAGATTGGAG 50
|||||
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeu 17
51 CTAGGCGCTGCGCTCGGATGGCTCTAGCCAGCATTTGAGCTAGCGGCTGC 100
|||||
17 rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeu 34
101 GCCCTGGCAGCGGTAGCCAGATTGAGCTAGCGCTGCGTCCGGGT 147
|||||
34 rgrprogllyserglyserclnlsprtsrtyrglyleuargprogly 49

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-878-748-4

seq_documentation_block:

; Sequence 4, Application US/08878748
; Patent No. 5969126

; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HOW P.A.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,748
; FILING DATE: 19-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-878-748-4

alignment_scores:
Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-306-689-10 x US-08-878-748-4 ..

Align seg 1/1 to: US-08-878-748-4 from: 1 to: 49

1 CAGATTGAGCTAGGCGCTGCGCTGGCAGCGGTTCTAAGATTGGAG 50
|||||
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeu 17
51 CTAGGCGCTGCGCTCGGATGGCTCTAGCCAGCATTTGAGCTAGCGGCTGC 100
|||||
17 rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeu 34
101 GCCCTGGCAGCGGTAGCCAGATTGAGCTAGCGCTGCGTCCGGGT 147
|||||
34 rgrprogllyserglyserclnlsprtsrtyrglyleuargprogly 49

seq_name: /cgn2_6/ptodata/2/1aa/6_COMB.pep:US-09-124-491-4

seq_documentation_block:

; Sequence 4, Application US/09124491
; Patent No. 6022960


```

:
: GENERAL INFORMATION:
: APPLICANT: POTTER, ANDREW A.
: APPLICANT: MANN, JOHN G.
: TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: REED & ROBINS LLP
: STREET: 285 HAMILTON AVENUE, SUITE 200
: CITY: PALO ALTO
: STATE: CA
: COUNTRY: USA
: ZIP: 94301
: COMPUTER READABLE FORM:
: MEDIUM TYPE: FLOPPY disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/124,491
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/694,865
: FILING DATE: 09-AUG-1996
: APPLICATION NUMBER: US 08/387,156
: FILING DATE: 10-FEB-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/960,932
: FILING DATE: 14-OCT-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/779,171
: FILING DATE: 16-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: MCCracken, THOMAS P.
: REGISTRATION NUMBER: 38,548
: REFERENCE/DOCKET NUMBER: 9001-0016.22
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415)327-3400
: TELEFAX: (415)327-3231
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 49 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-124-491-4

alignment_scores:
  Quality: 290.00      Length: 49
  Ratio: 5.918        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-09-124-491-4  ..
Align seg 1/1 to: US-09-124-491-4 from: 1 to: 49
1 CAGCATTTGAGCTACGGCTGGCCCTGGCAGCGGTCTCAAGATTGAG 50
|||||
1 GlnH1StrpSerTyrGlyLeuArpProGlySerGlySerGlnAspTrpSe 17
|||||
51 CTACGGCCCTGGCTCCGGGTGGCTCTAGCCAGCATTTGAGCTACGGCCCTGC 100
|||||
17 rTyGrGlyLeuArpProGlyGlySerSerGlnH1StrpSerTyrGlyLeuA 34
|||||
101 GCCCTGGCAGCGGTAGCCCAAGATTGAGCTAGCGCTGCGTCCGGGT 147
|||||
34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArpProGly 49

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-08-387-156-10
seq_documentation_block:

```

```

: Sequence 10, Application US/08387156
: Patent No. 5723129
:
: GENERAL INFORMATION:
: APPLICANT: POTTER, ANDREW A.
: APPLICANT: REDMOND, MARK J.
: APPLICANT: HUGHES, HOW P. A.
: TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
: NUMBER OF SEQUENCES: 28
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: REED & ROBINS
: STREET: 635 BRYANT STREET
: CITY: PALO ALTO
: STATE: CALIFORNIA
: COUNTRY: UNITED STATES OF AMERICA
: ZIP: 94301
: COMPUTER READABLE FORM:
: MEDIUM TYPE: FLOPPY disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/387,156
: FILING DATE: 10-FEB-1995
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/960,932
: FILING DATE: 14-OCT-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/779,171
: FILING DATE: 16-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: ROBINS, ROBERTA L.
: REGISTRATION NUMBER: 33,208
: REFERENCE/DOCKET NUMBER: 9001-0016.21
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 617-8999
: TELEFAX: (415) 327-3231
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 544 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-387-156-10

alignment_scores:
  Quality: 290.00      Length: 49
  Ratio: 5.918        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-08-387-156-10  ..
Align seg 1/1 to: US-08-387-156-10 from: 1 to: 544
1 CAGCATTTGAGCTACGGCTGGCCCTGGCAGCGGTCTCAAGATTGAG 50
|||||
494 GlnH1StrpSerTyrGlyLeuArpProGlySerGlySerGlnAspTrpSe 510
|||||
51 CTACGGCCCTGGCTCCGGGTGGCTCTAGCCAGCATTTGAGCTACGGCCCTGC 100
|||||
510 rTyGrGlyLeuArpProGlyGlySerSerGlnH1StrpSerTyrGlyLeuA 527
|||||
101 GCCCTGGCAGCGGTAGCCCAAGATTGAGCTAGCGCTGCGTCCGGGT 147
|||||
527 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArpProGly 542

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-694-865-10
seq_documentation_block:
: Sequence 10, Application US/08694865
: Patent No. 5837268

```

GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANNIS, JOHN G.
TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,865
FILING DATE: 09-AUG-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MCCRAKER, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-865-10

alignment_scores:
Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-08-694-865-10 ..
Align seg 1/1 to: US-08-694-865-10 from: 1. to: 544

1 CAGCATTGAGACTAGGCGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
494 GlnHistrpserTYrGlyLeuArgProGlySerGlySerGlnAspTyrPse 510
51 CTAGCGGCTGCGTCCGGGTGCTAGCCAGCATGGAGCTAGCGGCTGC 100
|||||
510 rTYrGlyLeuArgProGlySerGlySerGlnHistrpserTYrGlyLeuA 527
101 GCCCTGACAGCGGTAGCCAGATTGGAGCTAGCGGCTGCGGCT 147
|||||
527 rgrProGlySerGlySerGlnAspTyrPserTYrGlyLeuArgProGly 542

seq_name: /cgn2.6/ptodata/2/1aa/5B_COMB.pep:US-08-878-748-10

seq_documentation_block:
Sequence 10, Application US/08878748
Patent No. 5969126
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HOW P.A.
TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET

CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,748
FILING DATE: 19-JUN-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-748-10

alignment_scores:
Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-08-878-748-10 ..
Align seg 1/1 to: US-08-878-748-10 from: 1. to: 544

1 CAGCATTGAGACTAGGCGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
494 GlnHistrpserTYrGlyLeuArgProGlySerGlySerGlnAspTyrPse 510
51 CTAGCGGCTGCGTCCGGGTGCTAGCCAGCATGGAGCTAGCGGCTGC 100
|||||
510 rTYrGlyLeuArgProGlySerGlySerGlnHistrpserTYrGlyLeuA 527
101 GCCCTGACAGCGGTAGCCAGATTGGAGCTAGCGGCTGCGGCT 147
|||||
527 rgrProGlySerGlySerGlnAspTyrPserTYrGlyLeuArgProGly 542

seq_name: /cgn2.6/ptodata/2/1aa/6_COMB.pep:US-09-124-491-10

seq_documentation_block:
Sequence 10, Application US/09124491
Patent No. 6022960
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANNIS, JOHN G.
TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO

STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,491
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491-10

Alignment scores:
Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment block:
US-09-306-689-10 x US-09-124-491-10 ..
Align seg 1/1 to: US-09-124-491-10 from: 1 to: 544

1 CAGCATGTGAGCTACGGCCCTGCGCCCTGCGACGGTTCTCAAGATTGAG 50
|||||
494 GlnHisTrpSerTyrGlyLeuArGProGlySerGlnHisTrpSerTyrGlyLeuA 510
51 CTACGGCGCTGCGGTCCGGGTGCTAGCCAGCATTTGAGAGCTAGCGGCTCG 100
|||||
510 TyrGlyLeuArGProGlyGlySerSerGlnHisTrpSerTyrGlyLeuA 527
101 GCCCTGGCAGCGGTAGCCAGATTGAGAGCTAGCGGCTGCGCGGT 147
|||||
527 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArGProGly 542

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-694-865-16

seq_documentation_block:
Sequence 16, Application US/08694865
Patent No. 5837268
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANN, JOHN G.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP

STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,865
FILING DATE: 09-AUG-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 699 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-865-16

Alignment scores:
Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

Alignment block:
US-09-306-689-10 x US-08-694-865-16 ..
Align seg 1/1 to: US-08-694-865-16 from: 1 to: 699

1 CAGCATGTGAGCTACGGCCCTGCGCCCTGCGACGGTTCTCAAGATTGAG 50
|||||
9 GlnHisTrpSerTyrGlyLeuArGProGlySerGlnHisTrpSerTyrGlyLeuA 25
25 rTyrGlyLeuArGProGlyGlySerSerGlnHisTrpSerTyrGlyLeuA 42
51 CTACGGCGCTGCGGTCCGGGTGCTAGCCAGCATTTGAGAGCTAGCGGCTCG 100
|||||
42 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArGProGly 57
101 GCCCTGGCAGCGGTAGCCAGATTGAGAGCTAGCGGCTGCGCGGT 147
|||||

seq_name: /cgn2_6/ptodata/2/1aa/6_COMB.pep:US-09-124-491-16

seq_documentation_block:
Sequence 16, Application US/09124491
Patent No. 6022960
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANN, JOHN G.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,491
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, Thomas P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 699 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491-16

alignment_scores:
Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-09-124-491-16 ..

Align seq 1/1 to: US-09-124-491-16 from: 1 to: 699

```
1 CAGCATTGAGCTACGCGCTGCGCCCTGCGCAGCGGTTCTCAAGATTGAG 50
|||||
9 GlnHisTrpSerTyrGlyLeuArgProGlySerGlySerGlnAspTrpSe 25
51 CTAGGCGCTGCGTCCGGGTGCTCTAGCCAGCATTTGAGCTAGCGCCCTGC 100
|||||
25 rTyGclYleuArgProGlyGlySerSerGlnHisTrpSerTyrGlyLeuA 42
101 GCCCTGCGAGCGGTAGCCAGATTGAGCTACGCGCTGCGCGGT 147
|||||
42 rGrProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 57
```

seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-387-156-8

seq_documentation_block:

Sequence 8, Application US/08387156
Patent No. 5723129
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HUW P.A.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/387,156
FILING DATE: 10-FEB-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 977 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-387-156-8

alignment_scores:
Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-08-387-156-8 ..

Align seq 1/1 to: US-08-387-156-8 from: 1 to: 977

```
1 CAGCATTGAGCTACGCGCTGCGCCCTGCGCAGCGGTTCTCAAGATTGAG 50
|||||
927 GlnHisTrpSerTyrGlyLeuArgProGlySerGlySerGlnAspTrpSe 943
51 CTAGGCGCTGCGTCCGGGTGCTCTAGCCAGCATTTGAGCTAGCGCCCTGC 100
|||||
943 rTyGclYleuArgProGlyGlySerSerGlnHisTrpSerTyrGlyLeuA 960
101 GCCCTGCGAGCGGTAGCCAGATTGAGCTACGCGCTGCGCGGT 147
|||||
960 rGrProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 975
```

seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-694-865-8

seq_documentation_block:

Sequence 8, Application US/08694865
Patent No. 5837268
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANN, JOHN G.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,865
FILING DATE: 09-AUG-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 977 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-865-8

alignment_scores:
Quality: 290.00      Length: 49
Ratio: 5.918         Gaps: 0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-08-694-865-8 ..

Align seg 1/1 to: US-08-694-865-8 from: 1 to: 977

1 CAGCATTTGAGCTACGGCTCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
927 GlnH1StrpSerTyrGlyLeuArpProGlySerGlySerGlnAspTrpSe 943
943 rTyrGlyLeuArpProGlyGlySerSerGlnH1StrpSerTyrGlyLeuA 960
51 CACGAGCGCTCGCGGTGCTAGCCAGCATTTGAGCTAGCGGCTGC 100
|||||
943 rTyrGlyLeuArpProGlyGlySerSerGlnH1StrpSerTyrGlyLeuA 960
101 GCCCTGGAGCGGTAGCCAGATTGGAGCTAGCGGCTGCCTGGGT 147
|||||
960 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArpProGly 975

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-878-748-8

seq_documentation_block:
; Sequence 8, Application US/08878748
; Patent No. 5969126
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HOW P. A.
; TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,748
; FILING DATE: 19-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
```

```
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 977 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-748-8

alignment_scores:
Quality: 290.00      Length: 49
Ratio: 5.918         Gaps: 0
Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-08-878-748-8 ..

Align seg 1/1 to: US-08-878-748-8 from: 1 to: 977

1 CAGCATTTGAGCTACGGCTCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
927 GlnH1StrpSerTyrGlyLeuArpProGlySerGlySerGlnAspTrpSe 943
943 rTyrGlyLeuArpProGlyGlySerSerGlnH1StrpSerTyrGlyLeuA 960
51 CACGAGCGCTCGCGGTGCTAGCCAGCATTTGAGCTAGCGGCTGC 100
|||||
943 rTyrGlyLeuArpProGlyGlySerSerGlnH1StrpSerTyrGlyLeuA 960
101 GCCCTGGAGCGGTAGCCAGATTGGAGCTAGCGGCTGCCTGGGT 147
|||||
960 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArpProGly 975

seq_name: /cgn2_6/ptodata/2/1aa/6_COMB.pep:US-09-124-491-8

seq_documentation_block:
; Sequence 8, Application US/09124491
; Patent No. 6022960
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,491
; FILING DATE: 09-AUG-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE: 09-AUG-1996
; APPLICATION NUMBER: US 08/387,156
```

FILING DATE: 10-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 977 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491-8

alignment_scores:
Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-09-124-491-8 ..

Align seg 1/1 to: US-09-124-491-8 from: 1 to: 977

```
1 CAGATTGAGCTAGCGCCGCGCCGCGAGCGGTCTCAAGATTGAG 50
|||||
927 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnAspTrpSe 943
51 CTAGGCGCTGCGTCCGGGTGCTCTAGCCAGCATTTGAGCTAGCGCGTGC 100
|||||
943 rTyGlyLeuArgProGlySerSerGlnHisTrpSerTyrGlyLeuA 960
101 GCCCTGCGAGCGGTAGCCAGATTGAGCTAGCGCGTCCGGGT 147
|||||
960 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 975
```

seq_name: /cgn2_6/plodata/2/1aa/5A_COMB.pep:US-07-690-983D-45

seq_documentation_block:

Sequence 45, Application US/07690983D
Patent No. 5403586
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J.
APPLICANT: STEWART, Andrew G.
APPLICANT: TSONIS, Con G.
TITLE OF INVENTION: FUSION PROTEINS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/690,983D
FILING DATE: 25-JUN-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00373

FILING DATE: 24-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/148 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 44 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-690-983D-45

alignment_scores:
Quality: 187.50 Length: 49
Ratio: 4.934 Gaps: 3
Percent Similarity: 77.551 Percent Identity: 69.388

alignment_block:
US-09-306-689-10 x US-07-690-983D-45 ..

Align seg 1/1 to: US-07-690-983D-45 from: 1 to: 44

```
1 CAGATTGAGCTAGCGCCGCGCCCTGCGAGCGGTCTCAAGATTGAG 50
:::|||||
3 GlnHisTrpSerTyrGlyLeuArgProGly.....GlnHisTrpSe 16
51 CTAGGCGCTGCGTCCGGGTGCTCTAGCCAGCATTTGAGCTAGCGCGTGC 100
|||||
16 rTyGlyLeuArgProGly.....GlnHisTrpSerTyrGlyLeuA 30
101 GCCCTGCGAGCGGTAGCCAGATTGAGCTAGCGCGTCCGGGT 147
|||||
30 rGProGly.....GlnHisTrpSerTyrGlyLeuArgProGly 42
```

OM of: US-09-306-689-10 to: SPTREMBL_15:* out_format : pfs

Date: Mar 2, 2001 10:49 AM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frimex.n2p.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09306689/runat.02032001_102828.9616/app_query.fasta_1.2389
-DB=SPTREMBL_15 -QFMT=fastan -SUFFIX=rspt -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DEEXT=7.000 -START=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR=SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09306689.0CGN1_1.639 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-MAIT -THREADS=1

Search information block:

Query: US-09-306-689-10
Query length: 147
Database: SPTREMBL_15:*
Database sequences: 374700
Database length: 117207915
Search time (sec): 337.480000

score_list:

Sequence	Strd	Orig	ZScore	Escore	len	Documentation
sp_vertibrate:Q9PT52	+	95.50	166.91	0.0033	263	Q9PT52 agkistrodon halys blomh
sp_invertebrate:Q9WTY8	+	94.50	166.92	0.0037	836	Q9WTY8 rattus norvegicus (rat)
sp_invertebrate:Q9U617	+	81.00	153.65	0.1283	1729	Q9U617 drosophila melanogast
sp_invertebrate:Q9VT86	+	80.50	156.07	0.1564	1039	Q9VT86 drosophila melanogast
sp_plant:Q9SPF8	+	79.50	156.53	0.2144	715	Q9SPF8 arabidopsis thaliana (mc
sp_human:Q9U509	+	78.00	162.18	0.3769	197	Q9U509 homo sapiens (human). hy
sp_plant:Q9M815	+	76.00	148.69	0.5447	769	Q9M815 arabidopsis thaliana (mc
sp_invertebrate:PB2166	-	74.00	153.69	1.10	200	PB2166 locusta migratoria (m)
sp_mammal:Q97643	+	73.50	146.93	1.14	462	Q97643 lama glama (llama). fib
sp_human:Q13344	+	72.00	142.87	1.67	528	Q13344 homo sapiens (human). fu
sp_bacteria:Q9L070	+	72.00	140.93	1.62	701	Q9L070 streptomyces coelicolor
sp_vertibrate:Q57480	+	72.00	138.06	1.53	1068	Q57480 gallus gallus (chicken)
sp_mammal:Q97513	+	71.00	151.67	2.67	107	Q97513 macaca mulatta (rhesus m
sp_mammal:Q97655	+	71.00	151.23	2.65	114	Q97655 macaca mulatta (rhesus m
sp_plant:Q9X1L1	+	71.00	145.68	2.40	257	Q9X1L1 arabidopsis thaliana (mc
sp_plant:Q9SU12	+	71.00	138.34	2.10	754	Q9SU12 arabidopsis thaliana (mc
sp_plant:Q9A095	+	71.00	137.45	2.06	859	Q9A095 arabidopsis thaliana (mc
sp_invertebrate:Q22739	+	70.50	147.42	2.88	171	Q22739 caenorhabditis elegans
sp_human:Q16273	+	69.00	146.36	4.49	126	Q16273 homo sapiens (human). fu
sp_human:Q9P266	+	69.00	143.37	4.25	195	Q9P266 homo sapiens (human). fu
sp_invertebrate:Q96500	+	68.50	137.94	4.49	371	Q96500 blaberus craniifer. at
sp_plant:Q48709	+	68.50	133.55	4.14	705	Q48709 arabidopsis thaliana (mc
sp_human:Q9U5J4	+	68.00	144.26	5.87	126	Q9U5J4 homo sapiens (human). d
sp_plant:Q9U3L0	+	68.00	144.10	5.85	129	Q9U3L0 homo sapiens (human). hy
sp_invertebrate:Q9NB54	+	68.00	138.31	5.27	301	Q9NB54 arabidopsis thaliana (mc
sp_human:Q9UP1	+	68.00	135.88	5.04	430	Q9NB54 dreisena polymorpha (
sp_human:Q9U3L0	+	68.00	135.23	4.98	473	Q9U3L1 homo sapiens (human). ki
sp_vertibrate:Q13028	-	67.50	133.71	5.65	507	Q13028 homo sapiens (human). ki
sp_bacteria:Q33828	-	67.50	130.11	5.29	859	Q33828 thioacillus ferrooxidan
sp_vertibrate:Q9PRH0	-	67.00	144.38	8.00	91	Q9PRH0 anguilla japonica (japan
sp_plant:Q9X1H9	-	67.00	146.83	6.97	275	Q9X1H9 trifolium aestivum (wheat
sp_plant:Q9S155	+	66.50	134.90	7.85	313	Q9S155 arabidopsis thaliana (mc
sp_plant:Q9SLG9	+	66.50	128.74	7.01	772	Q9SLG9 arabidopsis thaliana (mc
sp_invertebrate:Q9VXK2	+	66.50	125.59	6.62	1224	Q9VXK2 drosophila melanogast
sp_invertebrate:Q9U9K2	+	66.00	131.73	8.63	427	Q9U9K2 aedes aegypti (yellow)
sp_invertebrate:Q22667	+	66.00	127.46	7.99	798	Q22667 caenorhabditis elegans
sp_invertebrate:Q24926	+	66.00	126.71	7.88	891	Q24926 eurytoma scolopes. per
sp_vertibrate:Q24925	+	66.00	126.69	7.88	894	Q24925 eurytoma scolopes. per
sp_vertibrate:Q9PVF5	+	66.00	124.64	7.59	1207	Q9PVF5 brachydanio rerio (zebr
sp_invertebrate:Q9NG18	+	65.00	131.52	11.70	324	Q9NG18 drosophila similans (d

sp_invertebrate:Q9VE67	+	65.00	117.04	8.98	2703	Q9VE67 drosophila melanog
sp_invertebrate:Q61603	+	65.00	117.01	8.97	2715	Q61603 drosophila melanog
sp_invertebrate:Q77072	+	64.50	129.31	13.10	384	Q77072 eisenia foetida (co
sp_invertebrate:Q9Z201	+	64.50	129.29	13.09	385	Q9Z201 cavia porcellus (guin
sp_human:Q9P2N8	+	64.50	124.79	12.06	745	Q9P2N8 homo sapiens (human)

seq_name: sp_vertibrate:Q9PT52

seq_documentation_block:

ID	Q9PT52	PRELIMINARY	PRT	263 AA.
AC	Q9PT52			
DT	01-MAY-2000 (TREMUREL_13, Created)			
DT	01-MAY-2000 (TREMUREL_13, Last sequence update)			
DT	01-JUN-2000 (TREMUREL_14, Last annotation update)			
DE	BPP-CNP PRECURSOR HOMOLOG.			
OS	Agkistrodon blomhoffi (Mamushi) (Gloydus blomhoffii).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;			
OC	Viperidae; Crotalinae; Agkistrodon.			
OX	NCBI_TaxID=61300;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=VENOM GLAND;			
RA	Murayama N.;			
RT	Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.			
RL	EMBL; AB020810; BAA36953.1;			
DR	INTERPRO; IPR000663;			
DR	PFAM; PF00212; ANP.1.			
DR	PRINTS; PR00710; NATPEPTIDES.			
DR	PROSITE; PS00263; NATRIURETIC_PEPTIDE.1.			
FT	CHAIN 31 41			BLOMOTIN.
FT	CHAIN 31 40			POTENTIATOR A.
FT	CHAIN 49 59			LEU3-BLOMOTIN.
FT	CHAIN 67 77			POTENTIATOR C.
FT	CHAIN 85 95			POTENTIATOR B.
FT	CHAIN 103 113			POTENTIATOR B.
FT	CHAIN 117 127			POTENTIATOR E.
FT	CHAIN 242 263			CNP-22.
SQ	SEQUENCE 263 AA; 27339 MW; 407BA9A572BF5FC8 CRC64;			

alignment_scores:

Quality:	95.50	Length:	63
Ratio: <td>3.183<td>Gaps:<td>3</td></td></td>	3.183 <td>Gaps:<td>3</td></td>	Gaps: <td>3</td>	3
Percent Similarity: <td>47.619<td>Percent Identity:<td>41.270</td></td></td>	47.619 <td>Percent Identity:<td>41.270</td></td>	Percent Identity: <td>41.270</td>	41.270

alignment_block:

US-09-306-689-10 x Q9PT52 ..

Align seg 1/1 to: Q9PT52 from: 1 to: 263

1	CAGCATTTGGAGCTACGCGCTGCGCCCTGGACGCGT.....	36
27	GLNGINTPserGLngIArgProProGLyProProIleProArgLeuVa	43
37	TCTCAAGATTGGAGCTACGCGCTGCGCGGTCCT.....	75
43	IValGLngINTPserGLngIleuProProGLyProProIleProArgL	60
76AGCAGCATTTGGAGCTACGCGCTGCGCCCTGGACGCGT.....	114
60	eUValValGLngINTPserGLngIleuProProGLyProProIlePro	76
115AGCAGCATTTGGAGCTACGCGCTGCGCTCGC	144
77	ProLeuValValGLngINTPserGLngIleuProPro	89

seq_name: sp_invertebrate:Q9WTY8

seq_documentation_block:

ID	Q9WTY8	PRELIMINARY	PRT	836 AA.
AC	Q9WTY8			

RA Nelson D.R., Nel

alignment_scores:
 Quality: 80.50 Length: 54
 Ratio: 2.439 Gaps: 4
 Percent Similarity: 61.111 Percent Identity: 40.741

alignment_block:
 US-09-306-689-10 x Q9VTR6 ..

Align seg 1/1 to: Q9VTR6 from: 1 to: 1039

seq_name: sp_plant:Q9SFB8

seq_documentation_block:
 ID Q9SFB8 PRELIMINARY: PRT: 715 AA.
 AC Q9SFB8;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE PUTATIVE PHOSPHATIDYLINOSITOL-4-PHOSPHATE 5-KINASE.
 GN F17A17.30.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
 RA Ronning C.M., Koo H., Fujii C.Y., Uterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.,
 RA "Arabidopsis thaliana chromosome III BAC F17A17 genomic sequence,"
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC013483; AAF21206.1;
 DR INTERPRO: IPR000169;
 DR INTERPRO: IPR002498;
 DR PFAM: PF01504; PfamK.1
 DR PROSITE: PS00639; THIOLEPROTEIN_HIS; UNKNOWN_1.
 KW Kinase.
 SQ SEQUENCE 715 AA; 81422 MW; 795D86CEB1FB52E9 CRC64;

alignment_scores:
 Quality: 80.50 Length: 54
 Ratio: 2.439 Gaps: 4
 Percent Similarity: 61.111 Percent Identity: 40.741

alignment_block:
 US-09-306-689-10 x Q9VTR6 ..

Align seg 1/1 to: Q9VTR6 from: 1 to: 1039

seq_name: sp_plant:Q9SFB8

seq_documentation_block:
 ID Q9SFB8 PRELIMINARY: PRT: 715 AA.
 AC Q9SFB8;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE PUTATIVE PHOSPHATIDYLINOSITOL-4-PHOSPHATE 5-KINASE.
 GN F17A17.30.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
 RA Ronning C.M., Koo H., Fujii C.Y., Uterback T.R., Barnstead M.E.,
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.,
 RA "Arabidopsis thaliana chromosome III BAC F17A17 genomic sequence,"
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC013483; AAF21206.1;
 DR INTERPRO: IPR000169;
 DR INTERPRO: IPR002498;
 DR PFAM: PF01504; PfamK.1
 DR PROSITE: PS00639; THIOLEPROTEIN_HIS; UNKNOWN_1.
 KW Kinase.
 SQ SEQUENCE 715 AA; 81422 MW; 795D86CEB1FB52E9 CRC64;

alignment_scores:
 Quality: 79.50 Length: 43
 Ratio: 3.180 Gaps: 1
 Percent Similarity: 58.140 Percent Identity: 41.860

alignment_block:
 US-09-306-689-10 x Q9SFB8 ..

Align seg 1/1 to: Q9SFB8 from: 1 to: 715

seq_name: sp_human:000509

seq_documentation_block:
 ID 000509 PRELIMINARY: PRT: 197 AA.
 AC 000509;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE HYPOTHETICAL 21.0 KDA PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA, AND PBMC;
 RX MEDLINE=96125193; PubMed=8543165;
 RA Gualandi G., Frezza D., Scotto A., Bianchi E., Gargano S.,
 RA Fruscalzo A., Calef E.;
 RT "Integration of an Epstein-Barr virus episome 3' into the gene
 RT encoding immunoglobulin heavy-chain alpha 1 in a lymphoblastoid cell
 RT line.";
 RL Gene 166:221-226(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA, AND PBMC;
 RA Gualandi G.;
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA, AND PBMC;
 RA Gualandi G.;
 RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA, AND PBMC;
 RA Frezza D.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA, AND PBMC;
 RA Frezza D.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X76785; CA54181.1;
 KW Hypothetical protein.
 SQ SEQUENCE 197 AA; 20977 MW; 7EF9725ED9D18CCF CRC64;

alignment_scores:
 Quality: 78.00 Length: 83

alignment_block:
 US-09-306-689-10 x Q9SFB8 ..

Align seg 1/1 to: Q9SFB8 from: 1 to: 715

seq_name: sp_human:000509

seq_documentation_block:
 ID 000509 PRELIMINARY: PRT: 197 AA.
 AC 000509;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE HYPOTHETICAL 21.0 KDA PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA, AND PBMC;
 RX MEDLINE=96125193; PubMed=8543165;
 RA Gualandi G., Frezza D., Scotto A., Bianchi E., Gargano S.,
 RA Fruscalzo A., Calef E.;
 RT "Integration of an Epstein-Barr virus episome 3' into the gene
 RT encoding immunoglobulin heavy-chain alpha 1 in a lymphoblastoid cell
 RT line.";
 RL Gene 166:221-226(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA, AND PBMC;
 RA Gualandi G.;
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA, AND PBMC;
 RA Gualandi G.;
 RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA, AND PBMC;
 RA Frezza D.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA, AND PBMC;
 RA Frezza D.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X76785; CA54181.1;
 KW Hypothetical protein.
 SQ SEQUENCE 197 AA; 20977 MW; 7EF9725ED9D18CCF CRC64;

Ratio: 2.438 Gaps: 3
Percent Similarity: 38.554 Percent Identity: 27.711

alignment_block:

US-09-306-689-10 x 000509 ..

Align seg 1/1 to: 000509 from: 1 to: 197

```

1 CAGCATTTGAGCTACGGCTGCGCTGCGAGCGGT..... 36
   ||| ||||| ..... |||||
61 GlnProtrpSerSerSerSerProglyProargHisSerAlaLeuAs 77
   .....TCTCAGATTGAGCTACGGCTGCGCTGCGAGCGGTCT... 75
   :||| :||| :||| :|||
77 pLeuValThrGlnProtrpSerSerSerSerProglyProargHis 94
   .....AGCCAGATTGAGCTACGGCTGCGCTGCGAGCGGTGC 108
   :||| :||| :||| :|||
94 eAlaLeuValLeuIleThrGlnProtrpThrSerSerLeuSerProgly 110
   ..... 114
109 AGCGGT..... 114
111 ProProHisSerAlaLeuAspLeuIleThrHisProglyProProHisLe 127
115 .....AGCCAGATTGAGCTACGGCTGCGCTGCGAGCGGT 147
   :||| :||| :||| :|||
127 uAlaLeuAspLeuIleThrGlnProtrpSerSerSerProgly 143

```

seq_name: sp_plant:09M815

seq_documentation_block:

```

ID 09M815 PRELIMINARY; PRT; 769 AA.
AC 09M815;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE PUTATIVE PHOSPHATIDYLINOSITOL-4-PHOSPHATE 5-KINASE.
GN 77Pl.4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLOMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Rönning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome I BAC T7P1 genomic sequence.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC018908; AAF27664.1; -.
KW kinase.
SQ SEQUENCE 769 AA; 87585 MW; 652P44182F493645 CRC64;

```

alignment_scores: Quality: 76.00 Length: 77
Ratio: 2.714 Gaps: 2
Percent Similarity: 36.364 Percent Identity: 29.870

alignment_block:

US-09-306-689-10 x 09M815 ..

Align seg 1/1 to: 09M815 from: 1 to: 769

```

7 TGAAGCTACGGCTGCGCTGCGAGCGGTCTCAGATTGAGCTACGG 56
   ||| ||||| ..... |||||
116 TrrpArgLuglInspglInspglSerLysSerLysTrpLysArg 132
   57 CCGCGCTCGCGGTGCTCT..... 75
   ||| ||||| ..... |||||
132 yAsnArgPheIleGlyAsnTrpLysLysGlyLysMetSerGlyArgGly 149

```

```

76 .....AGCCAGATTGAGCTACGGC 96
   :||| :|||
149 alMetSerTrpAlaAsnGlyAspLeuPheAsnGlyPheTrpLeuAsnGly 165
   ||||| |||||
97 CTGGCGCCCTGGCAGCGGTAGCCAA..... 120
   ||||| |||||
166 LeuArgHisGlySerGlyValTyrLysTyrAlaAspGlyGlyPheTyrPh 182
   .....GATTGAGCTACGGCTGCGCTGCGAGCGGT 147
   ||||| |||||
182 eGlyThrTrpSerArgGlyLeuLysAspGly 192

```

seq_name: sp_invertebrate:p82166

seq_documentation_block:

```

ID P82166 PRELIMINARY; PRT; 200 AA.
AC P82166;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
DE CUTICLE PROTEIN 19.8 (LMNCP19.8).
OS Locusta migratoria (migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridomorpha; Acridoidea; Acrididae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE.
RC TISSUE=FIFTH INSTAR LARVAE CUTICLE;
RA Anderson S.O.;
RL Submitted (NOV-1999) to the SWISS-PROT data bank.
CC -1- FUNCTION: COMPONENT OF THE CUTICLE OF MIGRATORY LOCUST WHICH
CC -1- CONTAINS MORE THAN 100 DIFFERENT STRUCTURAL PROTEINS.
CC -1- DOMAIN: THE TETRAPEPTIDE (A-A-P-(AV)) REPEATS FOUND THROUGHOUT THE
CC PROTECTIVE ENVELOPE OF OTHER SPECIES.
CC -1- SIMILARITY: CONTAINS A CUTICLE CONSENSUS DOMAIN.
DR INTERPRO: IPR000618; -.
DR PRAM: PR00379; Insect_cuticle; 1.
DR PRINTS: PR00947; CUTICLE.
DR PROSITE: PS00233; CUTICLE; 1.
KW Structural protein; Cuticle; Repeat.
FT DOMAIN 20 180
FT REPEAT 20 23 1.
FT REPEAT 26 29 2.
FT REPEAT 43 46 3.
FT REPEAT 126 129 4.
FT REPEAT 144 147 5.
FT REPEAT 150 153 6.
FT REPEAT 159 162 7.
FT REPEAT 177 180 8.
SQ SEQUENCE 200 AA; 19815 MW; 3802F2DB3BA6F92D CRC64;

```

alignment_scores: Quality: 74.00 Length: 47
Ratio: 2.387 Gaps: 2
Percent Similarity: 65.957 Percent Identity: 46.809

alignment_block:

US-09-306-689-10/rev x P82166 ..

Align seg 1/1 to: P82166 from: 1 to: 200

```

144 CGAAGCAGAGCGCTAGCTCAATCTGGCTACCGCT.....GCCAGGC 101
   ||||| ..... |||||
140 ArgThrValAlaAlaProAlaValAlaAlaAlaProValAlaArgAl 156
   100 GCAGGCTAGCTCAATGCTGGCTAGAGCCAGCGAGCGAGCGCTAG 51
   :||| :|||
156 aAlaIleAlaAlaProAlaTyrAla.....ThrTyrAlaAla 169

```


RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2):
RA Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.:
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL159139; CAB76891.1; -
KW Hypothetical protein.
SO SEQUENCE 701 AA; 73279 MW; 558DFCBCT7B412 CRC64;

alignment_scores:
Quality: 72.00 Length: 54
Ratio: 2.483 Gaps: 3
Percent Similarity: 53.704 Percent Identity: 38.889

alignment_block:
us-09-306-689-10 x Q9L0Z0 ..

Align seg 1/1 to: Q9L0Z0 from: 1 to: 701

```

13 TAGAGCGTGGCGCCGTCGAGCGGTTCTCAAGATTGAGCGCTGCG 62
   ::::: ||| ||| ||| ||| ::::: ||| ||
577 HIsGlyAlaAtgAspGlyHIsGlyArgGlyAspAspGlyAlaAr 593
   ::::: ||| ||| ||| ||| ::::: ||| ||
63 TCCG.....GGTGGCTGAGCGAGCATGAGC.... 90
   ||||| ||||| ||||| ||||| ||||| ||||
593 gProGlyHIsGlyThrGlyAspGlyGlyAspGlyHIsArgSerGluA 610
91 .....TAGCGCTGGCGCCGTCGAGCGGTTCAAGATTGAGCGCTGCG 135
   ||||| ||| ||| ||| ||||| ||| ::::: ||| ||
610 sPGLyTyrGlyAlaArgspGlyTyrGlyAlaArgsp...GLYtyrGly 625
136 CTGCGTCCGGGT 147
626 AlagInaspGly 629

```

seq_name: sp_vertebrate:057480

```

seq_documentation_block:
ID 057480 PRELIMINARY; PRT; 1068 AA.
AC 057480;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PHOTORECEPTOR GUANYLATE CYCLASE 1 (FRAGMENT).
GN GCL.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RHODE ISLAND RED; TISSUE-RETINA-PIGMENT EPITHELIUM-CHOROID;
RA Semple-Rowland S.L., Lee N.R., Van Hooser J.P., Palczewski K.,
RA Baehr W.:
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(1997).
DR EMBL: AF036942; AAC24500.1; -.
DR HSSP: Q02846; IAWL.
DR INTERPRO: IPR000719; -.
DR INTERPRO: IPR001054; -.
DR INTERPRO: IPR001245; -.
DR INTERPRO: IPR001828; -.
DR PFAM: PF00069; pkinase.1.
DR PFAM: PF00211; guanylate_cyc.1.
DR PFAM: PF01094; ANF_receptor.1.
DR PRINTS: PR00109; TYRKINASE.
DR PROSITE: PS00452; GUANYLATE_CYCLASES; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Lyase.

```

FT NON_TER 1 1
SO SEQUENCE 1068 AA; 118288 MW; 3E4C693E1597174B CRC64;

alignment_scores:
Quality: 72.00 Length: 41
Ratio: 3.130 Gaps: 2
Percent Similarity: 56.098 Percent Identity: 39.024

alignment_block:
us-09-306-689-10 x 057480 ..

Align seg 1/1 to: 057480 from: 1 to: 1068

```

28 GCGAGCGTCTCAAGATTGAGC...TAGCGCTGGCGTCCGGT.... 69
   ||||| ||||| ||||| ||||| ||||| |||||
356 GLySGlyAspArgLeuTyrProValTyrGlyLeuGluProGlyIleAr 372
70 .....GGCTGAGCGAGCATGAGCTAGCGCTGCGCC 103
   ||| ||| ||||| ||||| ||||| |||||
372 gGlyLeuAlaTyrArgGlyHIsSerValHIsTyrProHIsSerSer 389
104 CTGCGAGCGGTAGCCAGATTGG 126
   ||||| ||| ||| |||
389 roGlyThrAspSerGlyCysTyr 396

```

seq_name: sp_mammal:Q9TSI3

```

seq_documentation_block:
ID Q9TSI3 PRELIMINARY; PRT; 107 AA.
AC Q9TSI3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CHICKEN LUTEINIZING HORMONE-RELEASING HORMONE II (FRAGMENT).
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciae;
OC Cercopitheciae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Adler L.A., Sherwood N.M., Grendell R.L., Golos T.G., Terasawa E.:
RT "cDNA of a second form of luteinizing hormone releasing, chicken lHnR-
RT II, isolated from the non-human primate brain (abstract 632.8)".
RL Abstr. - Soc. Neurosci. 24:1607-1607(1998).
DR EMBL: AF104307; ABD13775.1; -.
DR INTERPRO: IPR002012; -.
DR PFAM: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
FT NON_TER 107 107
SO SEQUENCE 107 AA; 11823 MW; FACEB52703C3CB1D CRC64;

```

alignment_scores:
Quality: 71.00 Length: 39
Ratio: 2.731 Gaps: 2
Percent Similarity: 66.667 Percent Identity: 46.154

alignment_block:
us-09-306-689-10 x Q9TSI3 ..

Align seg 1/1 to: Q9TSI3 from: 1 to: 107

```

1 CAGATTGAGACTAGCGCTGCGCTGCG.....AGCGG 35
   ||||| ||||| ||||| ||||| ||||| |||||
25 GlnHIsTyrSerHIsGlyTyrTyrProGlyAlaGalaLeuSerSe 41
36 TTCGCAAGATTGAGCTAGCGCTGCG...CCGAGTGGCTGAGCGAGC 82
   ::::: ||||| ||| ::::: ||||| ||||| |||||
41 rAlaGlnaspProGlnaspAlaLeuArgProProAlaIleTyrProAlaG 58
83 ATTGAGCTAGCGGCTG 99

```

58 InlathrtyrGlyLeu 63

seq_name: sp_mammal:097655

seq_documentation_block:
ID 097655 PRELIMINARY; PRT; 114 AA.

AC 097655;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GONADOTROPIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN)
DE PRECURSOR.
GN GNRH2.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA White R.B., Urbanski H.F., Fernald R.D.;
RT "A second gene for gonadotropin-releasing hormone is expressed in the
RT rhesus macaque (Abstract #632.18)."
RL Abstr. - Soc. Neurosci. 24:1609-1609(1998).
CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC FOLLICLE-STIMULATING HORMONES.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC EMBL; AF097356; AAD09106.1; -.
DR INTERPRO: IPR002012; -.
DR PFM; PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
DR PRODOM; PD029787; -; 1.
KW Hormone; Amidation.
SQ SEQUENCE 114 AA; 12533 MW; 8B70D690D5BD5103 CRC64;

alignment_scores:
Quality: 71.00 Length: 39
Ratio: 2.731 Gaps: 2
Percent Similarity: 66.667 Percent Identity: 46.154

alignment_block:
US-09-306-689-10 x 097655 ..
Align seg 1/1 to: 097655 from: 1 to: 114

1 CAGCATTTGAGCTACGGCCTGCGCCCTGCG.....AGCG 35
|||||
25 GlnHISTPserHisGlyTrpTyrProGlyGlySArgAlaLeuSer 41
:::
36 TTCTCAAGATTGAGCTACGGCCTGCGT...CCGGGTGGCTCTAGCCAGC 82
:::
41 rAlaGlnAspProGlnAsnAlaLeuArgProAlaGlySerProAlaG 58
:::
83 ATTGAGCTACGGCCTG 99
58 InlathrtyrGlyLeu 63

seq_name: sp_plant:09X1L1

seq_documentation_block:
ID 09X1L1 PRELIMINARY; PRT; 257 AA.
AC 09X1L1;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE F19G14.22 PROTEIN.
GN F19G14.22
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Renning C.M., Benito M.,
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC F19G14 genomic sequence."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006438; AAD41986.1; -.
DR HSSP; P00303; 2CBP.
SQ SEQUENCE 257 AA; 27400 MW; DA59B7A93D53A17 CRC64;

alignment_scores:
Quality: 71.00 Length: 69
Ratio: 2.152 Gaps: 3
Percent Similarity: 47.826 Percent Identity: 26.087

alignment_block:

US-09-306-689-10 x 09X1L1 ..

Align seg 1/1 to: 09X1L1 from: 1 to: 257

7 TGGAGCTAGCGCCTGCGCCCT.....GGCAGCGGCTTC 38
|||||
87 TTPGlyTTPGlyGlyValProAsnAsnThrHisSerSerGlySerGly 103
:::
39 TCAAGATTG.....AGCTACGGCCTGCGCGGCG 70
:::
103 ySerGlyTTPGlyMetGlyProAsnAsnAsnTyrSerSerGlySerGly 120
:::
71 GCTTACGCCAGCATTTGAGCTACGGCCTGCGCCCTGCGAGC..... 111
|||||
120 ySerGlySerGlyTTPGlyTyrGlyHisSerTyrAsnAla 136
112GTAGCCAGATTGCGAG 128
137 ThrTyrAsnGlyProArgLysIleIleValGlyGlyAspLysGlyTTPH 153
129 CTACGCG 135
:::
153 TTYrGly 155

THIS PAGE BLANK (uspto)

FT PEPTIDE 6 15 GONADOLIBERIN I.
 FT PEPTIDE 19 >67 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT MOD_RES 6 6 ACTIVITY (BY SIMILARITY).
 FT MOD_RES 15 15 PYRROLIDONE CARBOXYLIC ACID (BY
 FT MOD_RES 15 15 SIMILARITY).
 FT AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
 FT NON_TER 67 67 SIMILARITY).
 FT SEQUENCE 67 AA: 7573 MW: 505394DAA261A3F2 CRC64;

alignment_scores:
 Quality: 76.50 Length: 19
 Ratio: 4.500 Gaps: 1
 Percent Similarity: 89.474 Percent Identity: 78.947

alignment_block:
 US-09-306-689-10 x GONL_MACMU ..

Align seg 1/1 to: GONL_MACMU from: 1 to: 67

67 GGTGGCTAGCCAGCATGAGCTAGCGCGCTGCGCGCTAG 116
 ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 2 GlycylserSerGlnHisTrpSerTyrGlyLeuArgProGly...Glyly 17

117 CCAGAT 123
 ::::|||
 17 sargasp 19

seq_name: SwissProt_39:GONL_HUMAN

seq_documentation_block:
 ID GONL_HUMAN STANDARD; PRT; 92 AA.

AC P01148;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
 DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
 DE HORMONE I) (GNRH I) (LULIBERIN I) (GONADORELIN); GNRH-ASSOCIATED
 DE PEPTIDE I].
 GN GNRH1 OR GNRH OR LHRH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-8936682; PubMed-2671939;
 RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
 RT "The complete nucleotide sequence of the human gonadotropin-releasing
 RT hormone gene."
 RL Nucleic Acids Res. 17:6403-6403(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-86094338; PubMed-2867548;
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting
 RT factor in human and rat."
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-85012739; PubMed-6090951;
 RA Seeburg P.H., Adelman J.P.;
 RT "Characterization of cDNA for precursor of human luteinizing hormone
 RT releasing hormone."
 RL Nature 311:666-668(1984).
 RN [4]
 RP SEQUENCE OF 24-33.
 RX MEDLINE-83126573; PubMed-6760865;
 RA Tan L., Rousseau P.;
 RT "The chemical identity of the immunoreactive LHRH-like peptide

RT biosynthesized in the human placenta."
 RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -1- PHARMACEUTICAL: AVAILABLE UNDER THE NAMES FACTREL (AYERST LABS),
 CC LUTREPUSE OR LUTRELFE (FERRING PHARMACEUTICALS) AND RELISORM
 CC (SERONO).
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: X01059; CAA25526.1; -;
 DR EMBL: M12578; AAA35916.1; -;
 DR EMBL: X15215; CAA33285.1; -;
 DR PIR: A01410; RHHUG.
 DR PIR: A26173; A26173.
 DR PIR: S05308; S05308.
 DR MIM: 152760; -;
 DR INTERPRO: IPR002012; -;
 DR PFAM: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Pharmaceutical; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 92 PROGNADOLIBERIN I.
 FT PEPTIDE 24 92 GONADOLIBERIN I.
 FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACT_SITE 26 26 ACTIVITY.
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 FT MOD_RES 16 16 W -> S (IN REF. 3).
 FT CONFLICT 16 16
 SQ SEQUENCE 92 AA: 10380 MW: 30A72221B076FA79 CRC64;

alignment_scores:
 Quality: 76.50 Length: 19
 Ratio: 4.500 Gaps: 1
 Percent Similarity: 89.474 Percent Identity: 78.947

alignment_block:
 US-09-306-689-10 x GONL_HUMAN ..

Align seg 1/1 to: GONL_HUMAN from: 1 to: 92

67 GGTGGCTAGCCAGCATGAGCTAGCGCGCTGCGCGCTAG 116
 ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 20 GlycylserSerGlnHisTrpSerTyrGlyLeuArgProGly...Glyly 35

117 CCAGAT 123
 ::::|||
 35 sargasp 37

seq_name: SwissProt_39:GONL_MOUSE

seq_documentation_block:
 ID GONL_MOUSE STANDARD; PRT; 90 AA.

AC P13562;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
 DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
 DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR
 DE I].
 GN GNRH1 OR GNRH.

AC P07490;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
 DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
 DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR
 DE I].
 GN GNRH1 OR GNRH.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-86094338; PubMed-2867548;
 RA Adelman J.P., Mason A.J., Haylick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting
 RT factor in human and rat.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89384661; PubMed-2476669;
 RA Bond C.T., Haylick J.S., Seeburg P.H., Adelman J.P.;
 RT "The rat gonadotropin-releasing hormone: SH locus: structure and
 RT hypothalamic expression.";
 RL Mol. Endocrinol. 3:1257-1262(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE-THYMUS;
 RA MEDLINE-93105480; PubMed-1468115;
 RA Maher C.G., Marchetti B., Leboucq R.D., Blalock J.E.;
 RT "Thymocytes express a mRNA that is identical to hypothalamic
 RT luteinizing hormone-releasing hormone mRNA.";
 RL Cell. Mol. Neurobiol. 12:447-454(1992).
 RN [4]
 RP SEQUENCE OF 1-47 FROM N.A.
 RC TISSUE-HEART;
 RX MEDLINE-87149087; PubMed-3547652;
 RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
 RT "Two mammalian genes transcribed from opposite strands of the same
 RT DNA locus.";
 RL Science 235:1514-1517(1987).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS: IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; S50870; AAB24572.1; -;
 DR EMBL; M12579; AAA41263.1; -;
 DR EMBL; M31670; AAA41264.1; -;
 DR EMBL; M15527; AAA42141.1; ALT_SEQ.
 DR EMBL; M15529; AAA42139.1; -;
 DR EMBL; M15528; -; NOT_ANNOTATED_CDS.
 DR PIR; B26173; RHRTG.
 DR PIR; A48410; A48410.
 DR INTERPRO; IPR002012; -;
 DR PFM; PF00446; GNRH; 1.
 DR PROSITE; PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 92 PROGNADOLIBERIN I.
 FT PEPTIDE 24 33 GONADOLIBERIN I.

FT PEPTIDE 37 92 PROLACTIN RELEASE-INHIBITING FACTOR I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT MOD_RES 24 24 ACTIVITY.
 FT MOD_RES 33 33 PYROLIDONE CARBOXYLIC ACID.
 SQ SEQUENCE 92 AA; 10500 MW; 494B564D8A3EB3 CRC64;
 alignment_scores:
 Quality: 74.00 Length: 14
 Ratio: 5.692 Gaps: 0
 Percent Similarity: 92.857 Percent Identity: 92.857
 alignment_block:
 US-09-306-689-10 x GON1_RAT ..
 Align seg 1/1 to: GON1_RAT from: 1 to: 92
 67 GGTGGCTGACGACGATGAGCTACGGCGCGCCGCGC 108
 ||| ||||||||||||||||||||||||||||||||
 20 GTCysSerSerGlnHisTrpSerTyrGlyLeuArgProGly 33.
 seq_name: SwissProt_39:SVS2_RAT
 seq_documentation_block:
 ID SVS2_RAT STANDARD; PRT; 414 AA.
 AC P22006;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE SEMINAL VESICLE SECRETORY PROTEIN II PRECURSOR (SVS II).
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE-90277684; PubMed-2351680;
 RA Harris S.E., Harris M.A., Johnson C.M., Bean M.F., Dodd J.G.,
 RA Marcus R.J., Carr S.A., Grubb J.W.;
 RT "Structural characterization of the rat seminal vesicle secretion II
 RT protein and gene.";
 RL J. Biol. Chem. 265:9896-9903(1990).
 CC -1- FUNCTION: THE RAT SEMINAL VESICLE CONTAINS SIX MAJOR ANDROGEN-
 CC DEPENDENT SECRETORY PROTEINS REFERRED TO AS SVS I-VI. THE SVS
 CC I-III PROTEINS APPEAR TO BE COMPONENTS OF THE RAT COPULATORY
 CC PLUG, WITH THE SVS II PROTEIN BEING THE MAJOR COMPONENT.
 CC -1- PTM: THE REPEATING UNIT APPEARS TO BE INVOLVED IN THE FORMATION OF
 CC THE COPULATORY PLUG VIA A TRANSGLUTAMINASE REACTION CROSS-LINKING
 CC GLUTAMINE AND LYSINE RESIDUES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J05443; AAA42192.1; -;
 DR PIR; A36443; A36443.
 DR INTERPRO; IPR002080; -;
 DR PROSITE; PS00515; SVS_II; 12.
 KW Testosterone; Seminal vesicle; Signal; Repeat; Copulatory plug.
 FT SIGNAL 1 22
 FT CHAIN 23 414
 FT MOD_RES 23 23 SEMINAL VESICLE SECRETORY PROTEIN II.
 FT DOMAIN 108 311 PYROLIDONE CARBOXYLIC ACID.
 FT REPEAT 108 120 13 X 13 AA TANDEM REPEATS.
 FT REPEAT 127 139 1.
 FT REPEAT 140 152 2.
 FT REPEAT 153 165 3.
 FT REPEAT 166 178 4.
 FT REPEAT 179 191 5.
 FT REPEAT 179 191 6.

```

FT REPEAT 192 204 7.
FT REPEAT 217 205 8.
FT REPEAT 224 236 9.
FT REPEAT 237 249 10.
FT REPEAT 257 269 11.
FT REPEAT 275 287 12.
FT REPEAT 299 311 13.
SQ SEQUENCE 414 AA: 45539 MW: C89E3FCE0C0EE92E CRC64:

```

```

alignment_scores:
  Quality: 71.00 Length: 42
  Ratio: 2.536 Gaps: 0
Percent Similarity: 66.667 Percent Identity: 42.857

```

alignment_block:
US-09-306-689-10 x SVS2_RAT ..

Align seq 1/1 to: SVS2_RAT from: 1 to: 414

```

10 AGCTACGCGCTGGCGCCCTGCTTCACAGATTGAGACTACGGCCT 59
   |||||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||
184 SerpHELyGlnMetLySerserGlySerGlnValLySerserPhgLygl 200
   :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||
60 GCGTCCGGGTGCTCTAGCCAGCATTTGAGCGCTGCGCCCTGGCA 109
   :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||
200 nMetLyAlaserGlnserGlnleLySerserPhgLyGlnArgLySersg 217
   :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||
110 GCGGTACGCAAGATTGAGACTACGGC 135
   :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||
217 lnglyGlyGlnlnglnserGly 225

```

seq_name: SwissProt_39:FUS_MOUSE

```

seq_documentation_block:
ID FUS_MOUSE STANDARD: PRT: 518 AA.
AC P56959;

```

DT 01-OCT-2000 (Rel. 40, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE RNA-BINDING PROTEIN FUS (PIGEPEN PROTEIN).

GN FUS.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

RN [1]

RP SEQUENCE FROM N.A.

RA Alapat S.R., Zhang M., Zhao X., Alliegro M.A., Alliegro M.C.,

RA Burdick C.A.;

RT "Regulation of p19en expression in mouse embryos.";

RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: BINDS BOTH SINGLE-STRANDED AND DOUBLE-STRANDED DNA AND PROMOTES ATP-INDEPENDENT ANNEALING OF COMPLEMENTARY SINGLE-STRANDED DNAs AND D-LOOP FORMATION IN SUPERHELICAL DOUBLE-STRANDED DNA. MAY PLAY A ROLE IN MAINTENANCE OF GENOMIC INTEGRITY (BY SIMILARITY).

CC -!- SUBUNIT: COMPONENT OF NUCLEAR RIBOPROTEIN COMPLEXES (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: AF224264; AAF70602.1; -

DR MGI:1353633; FUS.

DR PROSITE: PS50102; RRM; 1.

```

KM RNA-binding; DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;
KW Metal-binding.
FT DOMAIN 1 167 GLN/GLY/SER/TYR-RICH.
FT DOMAIN 168 265 GLY-RICH.
FT DOMAIN 278 364 RNA-BINDING (RRM).
FT DOMAIN 364 518 ARG/GLY-RICH.
FT ZN_FING 421 440 C4-TYPE (POTENTIAL).
SQ SEQUENCE 518 AA: 52673 MW: E06F231BFED78D6 CRC64:

```

```

alignment_scores:
  Quality: 70.00 Length: 49
  Ratio: 2.593 Gaps: 2
Percent Similarity: 55.102 Percent Identity: 40.816

```

alignment_block:
US-09-306-689-10 x FUS_MOUSE ..

Align seq 1/1 to: FUS_MOUSE from: 1 to: 518

```

13 TAGGCGCTGGCGCCCTGCGCGTTCACAGATTGAGACTACGGCCTGCG 62
   |||||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||
101 TyrlGlylnglnProAlaProserSerThSerGlyserTyrr..... 114
   :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||
63 TCCGGTGGCTCTAGCCAGCATTTGAGAC..... 90
   :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||
115 ...GlyGlyserSerSerSerSerTyrlGlylnglnProGlnserGly 130
   :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||
91 ..TAGGCGCTGGCGCCCTGCGCGTTCACAGATTGAGACTACGGC 135
   :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||
130 lYrGlyGlnGlnSerGlyTyrlGlylnglnGlnserTyrlGly 145

```

seq_name: SwissProt_39:FUS_BOVIN

```

seq_documentation_block:
ID FUS_BOVIN STANDARD: PRT: 512 AA.
AC Q28009;

```

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE RNA-BINDING PROTEIN FUS (PIGEPEN PROTEIN).

GN FUS.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;

OC Bovidae; Bovinae; Bos.

RN [1]

RP SEQUENCE FROM N.A.

RA TISSUE-AORTA;

RX MEDLINE-96175600; PubMed-8631501;

RA Alliegro M.C., Alliegro M.A.;

RT "A nuclear protein regulated during the transition from active to quiescent phenotype in cultured endothelial cells.";

RL Dev. Biol. 174:288-297(1996).

RN [2]

RP CARBOHYDRATE BINDING DOMAIN.

RA MEDLINE-20160719; PubMed-10694442;

RA Alliegro M.C.;

RT "A C-terminal carbohydrate-binding domain in the endothelial cell regulatory protein, p19en: new function for an EWS family member.";

RT Exp. Cell Res. 255:270-277(2000).

CC -!- FUNCTION: BINDS BOTH SINGLE-STRANDED AND DOUBLE-STRANDED DNA AND PROMOTES ATP-INDEPENDENT ANNEALING OF COMPLEMENTARY SINGLE-STRANDED DNAs AND D-LOOP FORMATION IN SUPERHELICAL DOUBLE-STRANDED DNA. MAY PLAY A ROLE IN MAINTENANCE OF GENOMIC INTEGRITY (BY SIMILARITY).

CC -!- SUBUNIT: COMPONENT OF NUCLEAR RIBOPROTEIN COMPLEXES (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: NUCLEAR, EXHIBITS DIFFUSE STAINING THROUGHOUT (EXCLUDING NUCLEOLI), TOGETHER WITH A SMALL NUMBER OF INTENSELY STAINED FOCAL POINTS, OR GRANULES, AND PUNCTATE STAINING ALONG THE NUCLEAR ENVELOPE.

CC -!- DOMAIN: THE C-TERMINAL DOMAIN BINDS CARBOHYDRATES.


```

FT DOMAIN 166 267 GLY-RICH.
FT DOMAIN 285 371 RNA-BINDING (RRM).
FT DOMAIN 371 526 ARG/GLY-RICH.
FT 2N FLNG 428 447 CA-TYPE (POTENTIAL).
FT SITE 266 267 BREAKPOINT FOR TRANSLLOCATION TO FORM
FT VARSPLIC 64 65 FUS/TLS-CHOP ONCOGENE.
FT CONFLICT 338 338 TG -> S (IN SHORT ISOFORM).
FT SEQUENCE 526 AA; 53426 MW; 88C8E263B7905549 CRC64;

```

```

alignment_scores:
  Quality: 69.00 Length: 45
  Ratio: 2.464 Gaps: 1
Percent Similarity: 62.222 Percent Identity: 40.000

```

alignment_block:
US-09-306-689-10 x FUS_HUMAN ..

Align seg 1/1 to: FUS_HUMAN from: 1 to: 526

```

13 TAGGGCGCTGGCGCTGAGCGGTTCAGATGAGCTAGCGGCTGGC 62
||||| ::::::::::: ::::| ::| ::| ::| ::| ::|
100 TyrlglnGlnProAlaProSerSerThrserylSerlylglySerse 116
:::||||| ::| ::| ::| ::| ::| ::|
63 TCCGGGTGGCTCTAGC.....CAGCATTTGAGCTAGCGGCTGGC 100
||||| ::| ::| ::| ::| ::| ::|
116 rserGlnSerSerSerlyrglyGlnProGlnSerGlySerlyrserGln 133
101 GCCCTGACGCGGTAGCCAGATGAGCTAGCGC 135
::| ::| ::| ::| ::| ::|
133 InProSerTyrglyGlyGlnGlnInSerlyrgly 144

```

seq_name: SwissProt_39:GONL_HAPBU

```

seq_documentation_block:
ID GONL_HAPBU STANDARD; PRT; 94 AA.
AC P51918; 093387;

```

```

DT 01-OCT-1996 (Rel. 34, Created)
DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 30-MAY-2000 (Rel. 39, last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE (LH-RH I) (LULIBERIN I).
GN GNRH1.
OS Haplochromis burtoni.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;
OC Cichlidae; Astatotilapia.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95396797; PubMed-7667296;
RA White S.A., Kasten T.L., Bond C.T., Adelman J.P., Fernald R.D.;
RT "Three gonadotropin-releasing hormone genes in one organism suggest
RT novel roles for an ancient peptide."
RL Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99061842; PubMed-9843638;
RA White R.B., Fernald R.D.;
RT "Ontogeny of gonadotropin-releasing hormone (GNRH) gene expression
RT reveals a distinct origin for GNRH-containing neurons in the
RT midbrain."
RL Gen. Comp. Endocrinol. 112:322-329(1998).
RN [3]
RP SEQUENCE OF 23-32.
RC TISSUE-PITUITARY;
RX MEDLINE-95372591; PubMed-7644702;
RA Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,
RA White S.A., Francis R.C., Fernald R.D., Licht P., Marby C.,
RA Sherwood N.M.;
RT Primary structure of solitary form of gonadotropin-releasing hormone
RT (GNRH) in cichlid pituitary: three forms of GNRH in brain of cichlid

```

```

RT and pumpkinseed fish."
RL Regul. Pept. 57:43-53(1995).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE
CC RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALMIC-PITUITARY-
CC GONADAL AXIS.
CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS
CC TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPOPHYSAL AXONS.
CC -1- MASS SPECTROMETRY: MW=1113.9; METHOD-MALDI; RANGE=23-32.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; 031865; AAC59691.1;
CC EMBL; AF076961; AAC27716.1;
CC INTERPRO: IPR002012;
CC PRAM: PR00446; GNRH; 1.
CC PROSITE: PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Signal; Multigene family.
CC FT SIGNAL 1 22
CC FT CHAIN 23 94 PROGONADOLIBERIN I.
CC FT PEPTIDE 23 32 GONADOLIBERIN I.
CC FT MOD_RES 36 94 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
CC FT MOD_RES 23 23 PYRROLIDONE CARBOXYLIC ACID.
CC FT MOD_RES 32 32 AMIDATION (G-33 PROVIDE AMIDE GROUP).
CC FT CONFLICT 86 94 ENGRHFRFK -> KMTGSRNRERFL (IN REF. 1).
CC SEQUENCE 94 AA; 10382 MW; E57DBA8333278D7 CRC64;

```

```

alignment_scores:
  Quality: 68.50 Length: 20
  Ratio: 4.567 Gaps: 1
Percent Similarity: 75.000 Percent Identity: 65.000

```

alignment_block:
US-09-306-689-10 x GONL_HAPBU ..

Align seg 1/1 to: GONL_HAPBU from: 1 to: 94

```

64 CCGGGTGGCTGTACCCAGATTGAGCTAGCGGCTGCGGCTGACGCG 113
||||| ::| ::| ::| ::| ::| ::|
18 ProGlnGlyCysGlnHisTrpSerTyrlglnSerProGly...G1 33
114 TAGCCAAGAT 123
||||| ::| ::| ::| ::| ::|
33 ylysaYgaSp 36

```

seq_name: SwissProt_39:SM34_LYTPI

```

seq_documentation_block:
ID SM34_LYTPI STANDARD; PRT; 335 AA.
AC 005904;

```

```

DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, last sequence update)
DT 15-DEC-1998 (Rel. 37, last annotation update)
DE 34 KDA SPICULE MATRIX PROTEIN PRECURSOR (LSM34).
OS Lytechinus pictus (Painted sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinozoa; Echinacea; Temnopneustidae;
OC Lytechinus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-GASTRULA;
RX MEDLINE-92077276; PubMed-1743395;
RA Livingston B.T., Shaw R., Bailey A., Wilt F.;
RA "Characterization of a cDNA encoding a protein involved in formation
RT of the skeleton during development of the sea urchin Lytechinus

```


RT isoamylase-hyperproducing mutant, Pseudomonas amyloclavata JD210.
RL Biochim. Biophys. Acta 1087:309-315(1990).
RN [3]
RP SEQUENCE OF 744-776 FROM N.A.
RC STRAIN-SB-15;
RX MEDLINE-89327147; PubMed-2753857;
RA Amenura A., Fujita M., Futai M.;
RT "transcription of the isoamylase gene (lam) in Pseudomonas
RL amyloclavata SB-15.";
RN J. Bacteriol. 171:4320-4325(1989).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE-98387895; PubMed-9719642;
RA Katsuya Y., Mezaki Y., Kubota M., Matsuura Y.;
RT "three-dimensional structure of Pseudomonas isoamylase at 2.2-A
RL resolution.";
RN J. Mol. Biol. 281:885-897(1998).
CC -I- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,6-ALPHA-D-GLUCOSIDIC BRANCH
CC LINKAGES IN GLYCOGEN, AMYLOPECTIN AND THEIR BETA-LIMITS DEXTRINS.
CC -I- INDUCTION: BY MALTOSE.
CC -I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY. ISOAMYLASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J03871; AAA25854.1; -
DR EMBL: X13378; CAA31754.1; -
DR PIR: A28109; A28109.
DR PDB: 1BR2; 12-AUG-98.
DR INTERPRO: IPR000461; -
DR PFAM: PF00128; alpha-amylase; 1.
DR KW Hydrolase; Glycosidase; Signal; 3D-structure.
FT SIGNAL 1 26
FT CHAIN 27 776
FT ACT_SITE 401 401
FT ACT_SITE 481 481
FT ACT_SITE 536 536
FT DISULFID 410 422
FT DISULFID 546 616
FT DISULFID 738 766
FT CONFLICT 8 8
FT CONFLICT 126 126
FT CONFLICT 169 171
FT CONFLICT 386 386
FT CONFLICT 413 416
FT CONFLICT 454 489
FT CONFLICT 555 556
FT CONFLICT 650 657
SQ SEQUENCE 776 AA: 83626 MW: F738BF8040246169 CRC64;
A -> G (IN REF. 1).
F -> C (IN REF. 1).
GAS -> AH (IN REF. 1).
L -> V (IN REF. 1).
GAVT -> AVH (IN REF. 1).
SGIDLEAFEPMAIGNSYLGFGPGGSEWNGLEFRDS ->
TWICLRNIGPSAATRTSWVDSRRYRVEMSVPRQ (IN
REF. 1).
WF -> S (IN REF. 1).
AFKRAHPA -> RSARHIP (IN REF. 1).

alignment_scores: quality: 66.00 Length: 37
Ratio: 2.750 Gaps: 1
Percent Similarity: 64.865 Percent Identity: 35.135

alignment_block:

US-09-306-689-10 x ISOA_PSEAV ..

Align seg 1/1 to: ISOA_PSEAV from: 1 to: 776

13 TACGGCTGGCGGCTGAGGCTTCACAGATGAGACTACGGCGCG 62
|||||:||||| ||| |||||
544 Tyserysasnglyalaasnasnserglalatrprotyrglyprose 560

63 TCCGGGTGCTAGCCAGCAT.....TGAGCTACGGCCTGGCCCTG 106
|||||:||||| |||:|
560 RASPGlyGlyThrsrThrsnTySerTrpAspGlnGlyMetSerAlaG 577
107 GCAGCGGTAGC 117
|||||:|||||
577 lyThrGlyAla 580


```

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-67 <RES>
A:Cross-references: GB:S75918; NID:g912831; PIDN:AA833096.1; PID:g912832
C:Superfamily: gonadolibrin

alignment_scores:
      Quality: 76.50      Length: 19
      Ratio: 4.500      Gaps: 1
      Percent Similarity: 89.474      Percent Identity: 78.947

alignment_block:
  US-09-306-689-10 x 178541 ..

  Align seg 1/1 to: 178541 from: 1 to: 67

      67 GGTGGCTTACCCAGCATTTGGAGCTACGGCTGCGCCCTTGCGACGGCTAG 116
          ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
      2 GtlyCyssterSerclnhtsrpSerTyrGlyLeuargProgly... Glyly 17

      117 CCAGAT 123
          ||| |||||
      17 sargasp 19

seq_name: p1rl.RHHUG

seq_documentation_block:
  gonadolibrin precursor - human
  N:Alternate names: gonadotropin releasing hormone (GNRH); luteinizing hormone releasing
  N:Comments: gonadolibrin-associated protein (GAP); progadolibrin
  C:Species: Homo sapiens (man)
  C:Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 18-Jun-1999
  C:Accession: S05308; A26173; A93342; A90108; A01410; S45718
  R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
  Nucleic Acids Res. 17, 6403-6404, 1989
  A:Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone ge
  A:Reference number: S05308; MUID:89366682
  A:Accession: S05308
  A:Status: translation not shown
  A:Molecule type: DNA
  A:Residues: 1-92 <NAV>
  A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956
  R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
  Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
  A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot
  A:Reference number: A94090; MUID:86094338
  A:Accession: A26173
  A:Molecule type: mRNA
  A:Residues: 1-92 <ADP>
  A:Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749
  A:Experimental source: hypothalamus
  R:Seeburg, P.H.; Adelman, J.P.
  Nature 311, 666-668, 1984
  A:Title: Characterization of cDNA for precursor of human luteinizing hormone releasing h
  A:Reference number: A93342; MUID:85012739
  A:Accession: A93342
  A:Molecule type: mRNA
  A:Residues: 1-15, 'S', 17-92 <SEE>
  A:Cross-references: GB:X01059; NID:g34356; PIDN:CAA25526.1; PID:g34357
  A:Experimental source: Placenta
  R:Tan, L.; Roussseau, P.
  Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982
  A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in
  A:Reference number: A90108; MUID:83126573
  A:Accession: A90108
  A:Molecule type: protein
  A:Residues: 24-33 <TAN>
  A:Experimental source: Placental trophoblasts
  R:Leibovitz, D.; Koch, F.; Pitzler, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amsterd
  EMBL Lett. 346, 203-206, 1994
  A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by th
  A:Reference number: S45718; MUID:94283597

```

```

A:Contents: annotation; degradation pathway of synthetic hormone
C:Genetics:
A:Gene: GDB:GNRH, LHRH, GRH
A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760
A:Map position: Bp21-8p11.2
A:Introns: 47/3; 79/3
C:Function:
A:Description: gonadoliberin stimulates pituitary secretion of luteotrophin and follitrop
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-92/Product: progonadoliberin #status predicted <PGN>
F:23-92/Product: gonadoliberin #status experimental <MAT>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

alignment_scores:
Quality: 76.50 Length: 19
Ratio: 4.500 Gaps: 1
Percent Similarity: 89.474 Percent Identity: 78.947

alignment_block:
US-09-306-689-10 x RHNG ..

Align seg 1/1 to: RHNG from: 1 to: 92

        67 GGTCCTAGCCACATTGAGACTAGCGCTGCCTGGCCGTGACGGTAG 116
           ||| | | | | | | | | | | | | | | | | | | | | |
          20 GLCYSSerSerGlnHisTrpserTyrGlyLeuATGProGly... GlyLy 35

        117 CCAGAT 123
           ::::|||
          35 sARGAsP 37

seq_name: p1r1:RHMSG

seq_documentation_block:
gonadoliberin precursor - mouse
N:Alternate names: gonadotropin-releasing hormone (GNRH); luteinizing hormone releasi
N:Contains: gonadoliberin; gonadoliberin-associated protein (GAP)
C:Species: Mus musculus (house mouse)
C>Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
C:Accession: A47578
R:Mason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikoli
Science 234, 1365-1371, 1986
A>Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible
A:Reference number: A47578; MUID:87069928
A:Accession: A47578
A:Molecule type: DNA
A:Residues: 1-90 <MAS>
A:Cross-references: EMBL:M14872; MID:g193576; PIDN:AAA37717.1; PID:g387175
C:Genetics:
A:Introns: 45/3; 77/3
C:Function:
A:Description: gonadoliberin stimulates pituitary secretion of lutropin and follitrop
A>Note: gonadoliberin-associated protein may have prolactin release inhibiting activi
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:22-31/Product: gonadoliberin #status predicted <GLB>
F:35-90/Product: gonadoliberin-associated protein #status predicted <GAP>
F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

alignment_scores:
Quality: 74.00 Length: 14
Ratio: 5.692 Gaps: 0
Percent Similarity: 92.857 Percent Identity: 92.857

```


Align seg 1/1 to: T42382 from: 1 to: 1068

```

28 GGACGGCGTCTCAAGATTGAGC...TAGCGCTGCGTCCGGCT..... 69
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
356 G1YLSGLYSPARGLYrarglYhISerValHIStrPrOHISerSerSer 372
   70 .....GCGTCTACGACGATTGGAGCTACGCGCTGGCC 103
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
372 gGlyLeuAlaTyrrarglYhISerValHIStrPrOHISerSerSer 389
   104 CTGGCAGCGGTAGCCCAAGATTGG 126
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
389 roglYThrAspSerGlyCysStrp 396

```

seq_name: p1r2:A36443

seq_documentation_block:

seminal vesicle secretory protein II precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 08-Mar-1991 #sequence_revision 08-Mar-1991 #text_change 05-Nov-1999
 C:Accession: A36443
 R:Harris, S.E.; Harris, M.A.; Johnson, C.M.; Bean, M.F.; Dodd, J.G.; Matusik, R.J.; Carr, J. Biol. Chem. 265, 9896-9903, 1990
 A:Title: Structural characterization of the rat seminal vesicle secretion II protein and
 A:Reference number: A36443; MUID:90277684
 A:Accession: A36443
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-414 <HAR>
 A:Cross-references: GB:J05443; NID:9207114; PIDN:AAA42192.1; PID:9207115

alignment_scores:
 Quality: 71.00 Length: 42
 Ratio: 2.536 Gaps: 0
 Percent Similarity: 66.667 Percent Identity: 42.857

alignment_block:
 US-09-306-689-10 x A36443 ..

Align seg 1/1 to: A36443 from: 1 to: 414

```

10 AGCTACGCGCTGCGCCCTGCGAGCGGTTCACAGATTGAGCTACGCGCT 59
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
184 SerheglYnMetlySerSerGlySerInVallySerPheglYgl 200
   .....GCGTCTACGACGATTGGAGCTACGCGCTGGCC 109
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
200 nMetlyAlaSerGlnSerGlnIlelySerPheglYGlnarGlySerG 217
   110 GCGGTAGCCCAAGATTGGAGCTACGCGC 135
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
217 lnglyglYglnleuGlnSerTyrgly 225

```

seq_name: p1r2:T16933

seq_documentation_block:

hypothetical protein T24D8.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T16933
 R:Martin, J.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid T24D8.
 A:Reference number: Z18609
 A:Accession: T16933
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-171 <HAR>
 A:Cross-references: EMBL:U04022; NID:g1065529; PID:g1065532; PIDN:AAA81446.1; CESP:T24D8
 C:Genetics:
 A:Gene: CESP:T24D8.5

A:Introns: 30/3; 91/2

alignment_scores:
 Quality: 70.50 Length: 47
 Ratio: 2.074 Gaps: 3
 Percent Similarity: 72.340 Percent Identity: 42.553

alignment_block:
 US-09-306-689-10 x T16933 ..

Align seg 1/1 to: T16933 from: 1 to: 171

```

16 GGCGTGGCGCCCTGCGACGCGTTCACAGATTGAGC...TAGCGCTGCG 62
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
105 GlyLeuArprroglYlYsArGserMeAlaTyrglYArGlnGlnlyPheAr 121
   63 TCCGGTGGCTCTAGCCAGCATTTGAGC...TAGCGCTGCGCCCTGGCA 109
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 gProglYlYsArGserMeAlaTyrglYArGlnGlnlyPheArproglYl 138
   110 GCGGTAGCCCAAGATTGAGC...TAGCGCTGCGTCCGGGT 147
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||
138 yArGserMeAlaTyrglYArGlnGlnlyPheArproglYl 151

```

seq_name: p1r4:S33798

seq_documentation_block:

FUS/CHOP mutant fusion protein - human
 N:Alternate names: TLS/CHOP mutant fusion protein
 C:Species: Homo sapiens (man)
 C:Date: 22-Nov-1993 #sequence_revision 30-Nov-1995 #text_change 20-Apr-2000
 C:Accession: S33798; S36158
 R:Crozat, A.; Aman, P.; Mandahl, N.; Ron, D.
 Nature 363, 640-644, 1993
 A:Title: Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma.
 A:Reference number: S33798; MUID:93288139
 A:Accession: S33798
 A:Molecule type: mRNA
 A:Residues: 1-462 <CRO>
 A:Cross-references: GB:S62138; NID:9386158; PIDN:AAB27103.1; PID:9386159
 R:Rabblits, T.H.; Forster, A.; Larson, R.; Nathan, P.
 Nature Genet. 4, 175-180, 1993
 A:Title: Fusion of the dominant negative transcription regulator CHOP with a novel ge
 A:Reference number: S36157
 A:Accession: S36158
 A:Molecule type: mRNA
 A:Residues: 1-462 <RAB>
 A:Cross-references: EMBL:X71427; NID:g939519; PIDN:CAA50558.1; PID:e1373214; PID:g4421
 C:Comment: This sequence is the chimeric product of a translocation mutation.
 C:Genetics:
 A:Gene: GADD153/FUS
 A:Map position: 12q43/16p11
 A:Note: TLS is a synonym for GDB:FUS
 C:Keywords: fusion protein

alignment_scores:
 Quality: 69.00 Length: 45
 Ratio: 2.464 Gaps: 1
 Percent Similarity: 62.222 Percent Identity: 40.000

alignment_block:
 US-09-306-689-10 x S33798 ..

Align seg 1/1 to: S33798 from: 1 to: 462

```

13 TAGCGCTGCGCCCTGCGACGCGTTCACAGATTGAGCTACGCGCTGCG 62
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
100 TyrlYglnglnPrroAlaProSerSerThSerlySerTyrglySerSe 116
   63 TCCGGTGGCTCTAGC.....CAGCATTTGAGCTACGCGCTGC 100
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:54:36 ; Search time 47.48 Seconds
(without alignments)
6.051 Million cell updates/sec

Title: US-09-306-689-5

Perfect score: 101

Sequence: 1 XHMSYGLRPPRPPPC 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	99.0	16	1	US-08-188-223-2
2	100	99.0	16	1	US-08-188-223-2
3	83.5	82.7	17	1	US-08-188-223-6
4	83.5	82.7	17	3	US-08-968-466-6
5	63	62.4	15	1	US-08-453-588-25
6	63	62.4	15	3	US-08-521-079-25
7	63	62.4	16	1	US-08-453-588-26
8	63	62.4	16	3	US-08-521-079-26
9	62	61.4	52	3	US-08-458-814-6
10	62	61.4	55	3	US-08-458-814-7
11	58	57.4	10	1	US-07-714-540-9
12	58	57.4	10	1	US-07-690-983D-2
13	58	57.4	10	1	US-07-690-983D-32
14	58	57.4	10	1	US-08-103-022-1
15	58	57.4	10	1	US-08-184-935-6
16	58	57.4	10	1	US-08-343-883-1
17	58	57.4	10	1	US-08-000-931-5
18	58	57.4	10	1	US-08-428-488-22
19	58	57.4	10	1	US-08-341-219-11
20	58	57.4	10	1	US-08-453-588-2
21	58	57.4	10	1	US-08-453-588-4
22	58	57.4	10	1	US-08-453-588-6
23	58	57.4	10	1	US-08-453-588-8
24	58	57.4	10	1	US-08-453-588-10
25	58	57.4	10	1	US-08-453-588-12
26	58	57.4	10	1	US-08-453-588-14
27	58	57.4	10	1	US-08-453-588-16
28	58	57.4	10	1	US-08-453-588-19

29	58	57.4	10	1	US-08-453-588-22	Sequence 22, Appl
30	58	57.4	10	1	US-08-188-223-3	Sequence 3, Appl
31	58	57.4	10	1	US-08-406-935-5	Sequence 1, Appl
32	58	57.4	10	1	US-08-591-917-1	Sequence 5, Appl
33	58	57.4	10	1	US-08-387-156-2	Sequence 2, Appl
34	58	57.4	10	1	US-08-474-555-1	Sequence 1, Appl
35	58	57.4	10	1	US-08-446-692-1	Sequence 1, Appl
36	58	57.4	10	1	US-08-242-678D-1	Sequence 1, Appl
37	58	57.4	10	2	US-08-796-598-6	Sequence 6, Appl
38	58	57.4	10	2	US-08-694-865-2	Sequence 2, Appl
39	58	57.4	10	2	US-08-694-865-18	Sequence 18, Appl
40	58	57.4	10	2	US-08-488-351A-1	Sequence 1, Appl
41	58	57.4	10	2	US-08-480-494B-1	Sequence 1, Appl
42	58	57.4	10	2	US-08-447-175A-6	Sequence 6, Appl
43	58	57.4	10	2	US-08-878-748-2	Sequence 2, Appl
44	58	57.4	10	3	US-08-521-079-2	Sequence 2, Appl
45	58	57.4	10	5	5492893-1	Patent No. 5492893

ALIGNMENTS

```

RESULT 1
US-08-188-223-2
; Sequence 2, Application US/08188223
; Patent No. 5686506
;
; GENERAL INFORMATION:
; APPLICANT: Grimes, Stephen
; APPLICANT: Sciablenki, Robert
; TITLE OF INVENTION: Immunogens Against Gonadotropin
; TITLE OF INVENTION: Releasing Hormone
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dimitrios T. Drivas, Esq.
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2787
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,223
; FILING DATE: 27-JAN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Drivas Esq., Dimitrios T.
; REGISTRATION NUMBER: 32,218
; REFERENCE/DOCKET NUMBER: 1102865-300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-819-8286
; TELEFAX: 212-354-8113
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: YES
;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..10
; OTHER INFORMATION: /note= "immunomimic"
;
; FEATURE:
; NAME/KEY: Region
; LOCATION: 11..16
; OTHER INFORMATION: /note= "spacer"
;
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1

```

OTHER INFORMATION: /label= pglu
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"
US-08-188-223-2

Query Match 99.0%; Score 100; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRGRPPPC 16
Db 2 HWSYGLRGRPPPC 16

RESULT 2
US-08-968-466-2
Sequence 2, Application US/08968466
Patent No. 613720
GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
APPLICANT: Scibienski, Robert
TITLE OF INVENTION: Immunogens Against Gonadotropin
TITLE OF INVENTION: Releasing Hormone
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dimitrios T. Drivas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 27-JAN-1994
APPLICATION NUMBER: US/08/968,466
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8286
TELEFAX: 212-354-8113
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
FEATURE:
NAME/KEY: Region
LOCATION: 1..10
OTHER INFORMATION: /note= "immunomimic"
FEATURE:
NAME/KEY: Region
LOCATION: 11..16
OTHER INFORMATION: /note= "spacer"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= pglu
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"
US-08-968-466-2

Query Match 99.0%; Score 100; DB 3; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRGRPPPC 16
Db 2 HWSYGLRGRPPPC 16

RESULT 3
US-08-188-223-6
Sequence 6, Application US/08188223
Patent No. 5688506
GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
APPLICANT: Scibienski, Robert
TITLE OF INVENTION: Immunogens Against Gonadotropin
TITLE OF INVENTION: Releasing Hormone
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dimitrios T. Drivas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 27-JAN-1994
APPLICATION NUMBER: US/08/188,223
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8286
TELEFAX: 212-354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= pglu
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"
FEATURE:
NAME/KEY: Region
LOCATION: 1..10
OTHER INFORMATION: /note= "immunomimic"
FEATURE:
NAME/KEY: Region
LOCATION: 11..17
OTHER INFORMATION: /note= "spacer"
US-08-188-223-6

Query Match 82.7%; Score 83.5; DB 1; Length 17;
Best Local Similarity 87.5%; Pred. No. 6.8e-05;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2 HWSYGLRPG-RPPPC 16
Db 2 HWSYGLRPGSSPPPC 17

RESULT 4
US-08-968-466-6

Sequence 6, Application US/08968466
Patent No. 6132720
GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
TITLE OF INVENTION: Immunogens Against Gonadotropin
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dimitrios T. Drivas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,466
FILING DATE: 27-JAN-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8286
TELEFAX: 212-354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= pglu
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"
FEATURE:
NAME/KEY: Region
LOCATION: 1..10
OTHER INFORMATION: /note= "immunomimic"
FEATURE:
NAME/KEY: Region
LOCATION: 11..17
OTHER INFORMATION: /note= "spacer"
US-08-968-466-6

Query Match 82.7%; Score 83.5; DB 3; Length 17;
Best Local Similarity 87.5%; Pred. No. 6.8e-05;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2 HWSYGLRPG-RPPPC 16
DB 2 HWSYGLRPGSSPPPC 17

RESULT 5
US-08-453-588-25
Sequence 25, Application US/08453588
Patent No. 5684145
GENERAL INFORMATION:
APPLICANT: Anna van der Zee, Irma Marianne van Die,
APPLICANT: Willem Pieter Martin Hoekstra,
APPLICANT: Josephus Theodorus Gielen,
TITLE OF INVENTION: Carrier system against GnRH
NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:
ADDRESSEE: Akzo No. 5684145el Patent Department
STREET: 1300 Piccard Drive, Suite 206
CITY: Rockville
STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,588
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/078,661
FILING DATE: 16-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Maly E. Gornley
REGISTRATION NUMBER: 34,409
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 258-5200
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-453-588-25

Query Match 62.4%; Score 63; DB 1; Length 15;
Best Local Similarity 90.9%; Pred. No. 0.023;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPGRP 12
DB 4 HWSYGLRPGDP 14

RESULT 6
US-08-521-079-25
Sequence 25, Application US/08521079
Patent No. 601983
GENERAL INFORMATION:
APPLICANT: Anna van der Zee, Irma Marianne van Die,
APPLICANT: Willem Pieter Martin Hoekstra,
APPLICANT: Josephus Theodorus Gielen,
TITLE OF INVENTION: Carrier system against GnRH
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Akzo Pharma
STREET: 1330 Piccard Drive
CITY: Rockville
STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/521,079
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/078,661
FILING DATE: 16-JUN-1993
APPLICATION NUMBER: EPA No. 601983 92.201.775.1
FILING DATE: 18-JUN-1992

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: William M. Blackstone
REGISTRATION NUMBER: 29,772
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 258-5200
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-521-079-25

Query Match 62.4%; Score 63; DB 3; Length 15;
Best Local Similarity 90.9%; Pred. No. 0.023;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRGRP 12
Db 4 HWSYGLRGRP 14

RESULT 7
US-08-453-588-26
Sequence 26, Application US/08453588
Patent No. 5684145
GENERAL INFORMATION:
APPLICANT: Anna van der Zee, Irma Marianne van Die,
APPLICANT: Willem Pieter Martin Hoekstra,
APPLICANT: Josephus Theodorus Gielen.
TITLE OF INVENTION: Carrier system against GNRH
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Akzo No. 5684145el Patent Department
STREET: 1300 Piccard Drive, Suite 206
CITY: Rockville
STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,588
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/078,661
FILING DATE: 16-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mary E. Gormley
REGISTRATION NUMBER: 34,409
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 258-5200
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-453-588-26

Query Match 62.4%; Score 63; DB 1; Length 16;
Best Local Similarity 90.9%; Pred. No. 0.024;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 HWSYGLRGRP 12

Db 5 HWSYGLRGRP 15

RESULT 8
US-08-521-079-26
Sequence 26, Application US/08521079
Patent No. 6019983
GENERAL INFORMATION:
APPLICANT: Anna van der Zee, Irma Marianne van Die,
APPLICANT: Willem Pieter Martin Hoekstra,
APPLICANT: Josephus Theodorus Gielen.
TITLE OF INVENTION: Carrier system against GNRH
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Akzo Pharma
STREET: 1330 Piccard Drive
CITY: Rockville
STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/521,079
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/078,661
FILING DATE: 16-JUN-1993
APPLICATION NUMBER: EPA NO. 6019983 92.201.775.1
FILING DATE: 18-JUN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: William M. Blackstone
REGISTRATION NUMBER: 29,772
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 258-5200
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-521-079-26

Query Match 62.4%; Score 63; DB 3; Length 16;
Best Local Similarity 90.9%; Pred. No. 0.024;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRGRP 12
Db 5 HWSYGLRGRP 15

RESULT 9
US-08-458-814-6
Sequence 6, Application US/08458814
Patent No. 6103243
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J
APPLICANT: DE AIZPURA, Henry J
APPLICANT: HOME, Peter
APPLICANT: RAND, Keith N
TITLE OF INVENTION: ORAL VACCINES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,814
FILING DATE: 02-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/327,822
FILING DATE: 18-OCT-1994
CLASSIFICATION: 424
PRIOR APPLICATION NUMBER: PCT/AU86/00135
FILING DATE: 14-MAY-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH3104
FILING DATE: 25-OCT-1985
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH0566
FILING DATE: 15-MAY-1985
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/155/BIANU
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 52 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-458-814-6

Query Match 61.4%; Score 62; DB 3; Length 52;
Best Local Similarity 83.3%; Pred. No. 0.096;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYGLRPGRP 13
DB 37 HWSYGLRPGDP 48

RESULT 10
US-08-458-814-7
Sequence 7, Application US/08458814
Patent No. 6103243
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J
APPLICANT: DE AIZPURA, Henry J
APPLICANT: HOME, Peter
APPLICANT: RAND, Keith N
TITLE OF INVENTION: ORAL VACCINES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,814
FILING DATE: 02-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/327,822
FILING DATE: 18-OCT-1994
CLASSIFICATION: 424
PRIOR APPLICATION NUMBER: PCT/AU86/00135
FILING DATE: 14-MAY-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH3104
FILING DATE: 25-OCT-1985
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH0566
FILING DATE: 15-MAY-1985
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/155/BIANU
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 55 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-458-814-7

Query Match 61.4%; Score 62; DB 3; Length 55;
Best Local Similarity 83.3%; Pred. No. 0.1;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYGLRPGRP 13
DB 39 HWSYGLRPGDP 50

RESULT 11
US-07-714-540-9
Sequence 9, Application US/07714540
Patent No. 5262521
GENERAL INFORMATION:
APPLICANT: Almquist, Ronald G.
APPLICANT: Toll, Lawrence
TITLE OF INVENTION: ISOLATED ATRIAL PEPTIDE-DEGRADING
TITLE OF INVENTION: ENZYME AND NOVEL COMPOUNDS USEFUL AS INHIBITORS THEREOF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Itell & Manella
STREET: 545 Middlefield Road, Suite 200
CITY: Menlo Park
STATE: California
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/714,540
FILING DATE: 19910607
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Dianne E.

REGISTRATION NUMBER: 31,292
REFERENCE/DOCKET NUMBER: 8500-0135.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-327-7250
TELEFAX: 415-327-2951
TELEX: 706141
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-714-540-9

Query Match 57.4%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
Db 2 HWSYGLRPG 10

RESULT 12
US-07-690-983D-2
Sequence 2, Application US/07690983D
Patent No. 5403586
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J.
APPLICANT: STEWART, Andrew G.
APPLICANT: TSONIS, Con G.
TITLE OF INVENTION: FUSION PROTEINS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/690,983D
FILING DATE: 25-JUN-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00373
FILING DATE: 24-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/148 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-07-690-983D-2

Query Match 57.4%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
Db 2 HWSYGLRPG 10

RESULT 13
US-07-690-983D-32
Sequence 32, Application US/07690983D
Patent No. 5403586
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J.
APPLICANT: STEWART, Andrew G.
APPLICANT: TSONIS, Con G.
TITLE OF INVENTION: FUSION PROTEINS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/690,983D
FILING DATE: 25-JUN-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00373
FILING DATE: 24-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/148 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-690-983D-32

Query Match 57.4%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
Db 2 HWSYGLRPG 10

RESULT 14
US-08-103-022-1
Sequence 1, Application US/08103022
Patent No. 5413990
GENERAL INFORMATION:
APPLICANT: Haviv, Fortuna
APPLICANT: Fitzpatrick, Timothy D.
APPLICANT: Swenson, Rolf E.
APPLICANT: Nichols, Charles J.
APPLICANT: Mott, Nicholas A.
TITLE OF INVENTION: N-Terminus Modified Analogs of LHRH
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edward H. Gorman, Jr., Dept. 377
STREET: Abbott Laboratories, One Abbott Park Road

CITY: NO. 5413990Lh Chicago
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/103,022
FILING DATE: 05-OCT-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Janssen, Jerry F.
REGISTRATION NUMBER: 29,175
REFERENCE/DOCKET NUMBER: 5389, US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 938-7742
TELEFAX: (708) 938-2623
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "Xaa at position 1 is a
5-oxo-prolyl aminoacyl residue."
US-08-103-022-1

Query Match 57.4%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
DB 2 HWSYGLRPG 10

RESULT 15
US-08-184-935-6
Sequence 6, Application US/08184935
Patent No. 5476770
GENERAL INFORMATION:
APPLICANT: PRADELLES, PHILIPPE
TITLE OF INVENTION: IMMUNOMETRIC DETERMINATION OF AN ANTIGEN
TITLE OF INVENTION: OR HAPTEN
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: P.C. JEFFERSON DAVIS HIGHWAY, SUITE 400
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/184,935
FILING DATE: 24-JAN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NO. 547670man F.
REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 846-286-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 10
OTHER INFORMATION: /note= "C-terminal amide"
US-08-184-935-6

Query Match 57.4%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||||
DB 2 HWSYGLRPG 10

Search completed: March 2, 2001, 10:54:36
Job time: 601 sec

THIS PAGE BLANK (uspro)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: March 2, 2001, 11:00:48 ; Search time 97.13 Seconds
(without alignments)
19.307 Million cell updates/sec

Title: US-09-306-689-5
Perfect score: 101
Sequence: 1 XHMYGLRPGRRPPPC 16

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_15:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.plant:*
11: sp.todent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	57.4	91	13	Q9PRH0
2	55	54.5	969	5	Q17795
3	54	53.5	381	5	Q9VWD2
4	53	52.5	112	13	Q88020
5	51.5	51.5	87	12	Q9Y126
6	52	51.5	95	13	Q73812
7	52	51.5	99	13	Q91A10
8	52	51.5	112	12	Q04273
9	52	51.5	112	12	Q88050
10	52	51.5	112	12	Q88057
11	52	51.5	112	12	Q88058
12	52	51.5	112	12	Q88064
13	52	51.5	112	12	Q88068
14	52	51.5	112	12	Q88125
15	52	51.5	112	12	Q88131
16	52	51.5	112	12	Q88137
17	52	51.5	112	12	Q88145
18	52	51.5	112	12	Q89827
19	52	51.5	112	12	Q07393

20	52	51.5	112	12	Q85604	085604 human t-cell
21	52	51.5	112	12	Q89552	089552 chimpanzee
22	51	50.5	263	13	Q9PT52	09PT52 agkistrodon
23	51	50.5	2061	5	Q9VUH9	09VUH9 drosophila
24	50	49.5	90	13	Q91A02	091A02 rana dybows
25	50	49.5	112	12	Q76623	076623 human immun
26	50	49.5	112	12	P88144	088144 human immun
27	50	49.5	112	12	Q90159	090159 human immun
28	50	49.5	112	12	Q90158	090158 human immun
29	50	49.5	112	12	Q90157	090157 human immun
30	50	49.5	112	12	Q90156	090156 human immun
31	50	49.5	112	12	Q9PX03	09PX03 human immun
32	50	49.5	211	5	045098	045098 caenorhabdi
33	50	49.5	480	6	Q06018	006018 ovis aries
34	49.5	48.0	423	12	Q12272	012272 simian t-ce
35	49	48.5	373	4	Q9NP08	09NP08 homo sapien
36	49	48.5	645	4	Q9P282	09P282 homo sapien
37	49	48.5	754	5	Q9NMW3	09NMW3 caenorhabdi
38	49	48.5	2129	5	Q9Q1R8	09Q1R8 caenorhabdi
39	48.5	48.0	1049	4	Q94957	094957 homo sapien
40	48	47.5	329	12	Q09798	009798 human herpe
41	48	47.5	598	3	Q9P6X0	09P6X0 neurospora
42	48	47.5	629	5	Q07997	007997 chironomus
43	48	47.5	635	12	Q69087	069087 human herpe
44	48	47.5	1051	5	Q9Y150	09Y150 drosophila
45	48	47.5	1579	5	Q9Y4S1	09Y4S1 drosophila

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	91 AA.
Q9PRH0	Q9PRH0			
AC	Q9PRH0			
DT	01-MAY-2000 (TREMBlrel. 13, Created)			
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)			
DT	01-JUN-2000 (TREMBlrel. 14, Last annotation update)			
DE	PREPRO-MGNRH PRECURSOR.			
OS	Anguilla japonica (Japanese eel).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;			
OC	Anguillidae; Anguilla.			
OX	NCBI_TaxID=7937;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-BRAIN:			
RA	Kubo K., Suetake H., Aida K.;			
RT	"Expression of two gonadotropin-releasing hormone (GNRH) precursor			
RT	genes in various tissues of the Japanese eel and evolution of GNRH.";			
RL	Zool. Sci. 16:471-478(1999).			
RL	(2)			
RP	SEQUENCE FROM N.A.			
RA	Kubo K., Suetake H., Aida K.;			
RT	"A splicing variant for the prepro-mammalian gonadotropin-releasing			
RT	hormone (prepro-mGNRH) mRNA is present in the brain and various			
RT	peripheral tissues of the Japanese eel.";			
RL	Zool. Sci. 16:645-651(1999).			
DR	EMBL; AB026989; BAA82608.1;			
DR	EMBL; AB026991; BAA83597.1;			
DR	INTERPRO: IPR020122;			
DR	PFAM; PF00446; GNRH; 1.			
DR	PROSITE; PS00473; GNRH; 1.			
KW	Signal.			
FT	SIGNAL.			
FT	CHAIN			
FT	CHAIN			
SO	SEQUENCE			

Query Match 57.4%; Score 58; DB 13; Length 91;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
1111111111
Db 24 HWSYGLRPG 32

RESULT 2
ID 017795 PRELIMINARY; PRT: 969 AA.
AC 017795;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE SIMILARITY TO WISCOTT-ALDRICH SYNDROME PROTEINS.
GN CG14234.4.
OS Eukaryota: Metazoa: Nematoda: Chromadorea: Rhabditida: Rhabditoidea:
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Keshaw J., Kirsten T., Laister L., Latreille P.,
RA Lighting J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Hawkins J.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U58751; AAB00657.1; -;
DR INTERPRO: IPR000095; -;
DR INTERPRO: IPR000697; -;
DR INTERPRO: IPR001960; -;
DR INTERPRO: IPR003124; -;
DR PFAM: PF00568; WH1; 1.
DR PFAM: PF00786; PBD; 1.
DR PFAM: PF02205; WH2; 2.
SQ SEQUENCE 969 AA; 105790 MW; BE2D115D0C620BE8 CRC64;

Query Match 54.5%; Score 55; DB 5; Length 969;
Best Local Similarity 75.0%; Pred. No. 7;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 SYGLRPGRPPPP 15
1111111111
Db 727 SHGLAPARPPPP 738

RESULT 3
ID 09VMD2 PRELIMINARY; PRT: 381 AA.
AC 09VMD2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)

DE CG14234 PROTEIN.
GN CG14234.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:
OC Pterygota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha:
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Plannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL: AE003513; AAF49014.1; -;
DR FLYBASE: FBgn0031065; CG14234.
SQ SEQUENCE 381 AA; 41732 MW; 77317DA4EC40A863 CRC64;

Query Match 53.5%; Score 54; DB 5; Length 381;
Best Local Similarity 53.3%; Pred. No. 4;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGRPPPP 16
1111111111
Db 205 NWIGMRHPSPPPC 219

RESULT 4
ID 088020 PRELIMINARY; PRT: 112 AA.
AC 088020;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE (M316ZA) DNA, COMPLETE CDS'S.
GN VPX.
OS Chimpanzee immunodeficiency virus (SIVcpz) (CIV).

OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID-11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94016835; PubMed-8411355;
 RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;
 RT "Analysis of simian immunodeficiency virus sequence variation in
 RT tissues of rhesus macaques with simian AIDS."
 RL J. Virol. 67:6522-6534(1993).
 DR EMBL: L22807; AAA47688.1; -.
 DR INTERPRO: IPR000012; -.
 DR PFAM: PF00522; VPR; 1.
 DR PRINTS: PRO0444; HIVPRVFX.
 SQ SEQUENCE 112 AA; 12945 MW; A933BF982AFD7BE8 CRC64;

Query Match 52.5%; Score 53; DB 12; Length 112;
 Best Local Similarity 64.3%; Pred. No. 1.7;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 HMSGSLRPGPPP 15
 Db 94 HGAEGWRPGPPPP 107

RESULT 5
 OY126 PRELIMINARY; PRT; 87 AA.
 AC 09Y126;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, last annotation update)
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN)
 DE (FRAGMENT)
 OS Sparus aurata (Gilthead sea bream).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 OC Sparidae; Sparus.
 OX NCBI_TaxID-8175;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RA Nabissi M.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC FOLLICLE-STIMULATING HORMONES.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC EMBL: AF046801; AAD02427.1; -.
 DR INTERPRO: IPR002012; -.
 DR PFAM: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation.
 FT NON_TER 1
 FT NON_TER 87
 SQ SEQUENCE 87 AA; 9871 MW; 0D2463533D96782A CRC64;

Query Match 51.5%; Score 52; DB 13; Length 87;
 Best Local Similarity 88.9%; Pred. No. 1.9;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HMSGSLRPG 10
 Db 22 HMSGSLSPG 30

RESULT 6
 OY3812 PRELIMINARY; PRT; 95 AA.
 AC 073812;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN).
 OS Morone saxatilis (Striped bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 OC Moronidae; Morone.
 OX NCBI_TaxID-34816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chow M.M., Kight K.E., Gotthelf Y., Alok D., Zohar Y.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC FOLLICLE-STIMULATING HORMONES.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC EMBL: AF056314; AAD03817.1; -.
 DR INTERPRO: IPR002012; -.
 DR PFAM: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 DR PRODOM: PD005581; -; 1.
 KW Hormone; Amidation.
 SQ SEQUENCE 95 AA; 10411 MW; 980C6988FC279BFC CRC64;

Query Match 51.5%; Score 52; DB 13; Length 95;
 Best Local Similarity 88.9%; Pred. No. 2.1;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HMSGSLRPG 10
 Db 24 HMSGSLSPG 32

RESULT 7
 OY1A10 PRELIMINARY; PRT; 99 AA.
 AC 09Y1A10;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)
 DE GONADOTROPIN-RELEASING HORMONE SEABREAM ISOFORM.
 OS Dicentrarchus labrax (European sea bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 OC Moronidae; Dicentrarchus.
 OX NCBI_TaxID-13489;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RA Gonzalez-Martinez D., Madigou T., Zmora N., Anglade I., Zanny S.,
 RA Zohar Y., Elizur A., Munoz-Cueto J.A., Kah O.;
 RT "Differential expression of three different prepro-GNRH
 RT (Gonadotrophin-releasing hormone) messengers in the brain of the
 RT European sea bass (Dicentrarchus labrax)."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RA Zmora N., Zohar Y., Elizur A.;
 RT "3 GNRH form in the seabass Dicentrarchus labrax";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF224279; AAF62898.1; -.
 SQ SEQUENCE 99 AA; 10758 MW; ECA8AEC93CC02904 CRC64;

Query Match 51.5%; Score 52; DB 13; Length 99;
 Best Local Similarity 88.9%; Pred. No. 2.1;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HMSGSLRPG 10
 Db 28 HMSGSLSPG 36

RESULT 8
ID 004273 PRELIMINARY; PRT; 112 AA.
AC 004273;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE VPX PROTEIN (X ORF PROTEIN).
GN VPX.
OS Chimpanzee immunodeficiency virus (SIV/cpz) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RA Benveniste R.E., Heidecker G., Greenwood J., Gonda M.A.;
RL Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.
CC -1- MISCBLANEOUS: THIS IS A CAPTIVE MACAQUE NEMESTRINA ISOLATE.
DR EMBL: M32741; AAA91933.1; -
DR INTERPRO: IPR000012; -
DR PFAM: PF00522; VPR; 1.
DR PRINTS: PR00444; HIVPRVXP.
KM AIDS.
SQ SEQUENCE 112 AA; 12904 MW; 7E1564FA5AFD7BEA CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;
Best Local Similarity 64.3%; Pred. No. 2.4;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 HWSYGLRGRGPPPP 15
| : | | | | | | |
Db 94 HGAGGWRGPPPP 107

RESULT 9
ID 088050 PRELIMINARY; PRT; 112 AA.
AC 088050;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE (MM316X) DNA, COMPLETE CDS'S.
GN VPX.
OS Chimpanzee immunodeficiency virus (SIV/cpz) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94016835; PubMed=8411355;
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;
RT "Analysis of simian immunodeficiency virus sequence variation in
RT tissues of rhesus macaques with simian AIDS."
RL J. Virol. 67:6522-6534(1993).
DR EMBL: L22804; AAA47679.1; -
DR INTERPRO: IPR000012; -
DR PFAM: PF00522; VPR; 1.
DR PRINTS: PR00444; HIVPRVXP.
SQ SEQUENCE 112 AA; 12872 MW; 44BD094EF1955B60 CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;
Best Local Similarity 64.3%; Pred. No. 2.4;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 HWSYGLRGRGPPPP 15
| : | | | | | | |
Db 94 HGAGGWRGPPPP 107

RESULT 10
OY 2 HWSYGLRGRGPPPP 15
| : | | | | | | |
Db 94 HGAGGWRGPPPP 107

ID 088057 PRELIMINARY; PRT; 112 AA.
AC 088057;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE (MM316K) DNA, COMPLETE CDS'S.
GN VPX.
OS Chimpanzee immunodeficiency virus (SIV/cpz) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94016835; PubMed=8411355;
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;
RT "Analysis of simian immunodeficiency virus sequence variation in
RT tissues of rhesus macaques with simian AIDS."
RL J. Virol. 67:6522-6534(1993).
DR EMBL: L22791; AAA72399.1; -
DR INTERPRO: IPR000012; -
DR PFAM: PF00522; VPR; 1.
DR PRINTS: PR00444; HIVPRVXP.
SQ SEQUENCE 112 AA; 12842 MW; BDB78F982AFD7BF2 CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;
Best Local Similarity 64.3%; Pred. No. 2.4;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 HWSYGLRGRGPPPP 15
| : | | | | | | |
Db 94 HGAGGWRGPPPP 107

RESULT 11
ID 088058 PRELIMINARY; PRT; 112 AA.
AC 088058;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE (MM316L) DNA, COMPLETE CDS'S.
GN VPX.
OS Chimpanzee immunodeficiency virus (SIV/cpz) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94016835; PubMed=8411355;
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;
RT "Analysis of simian immunodeficiency virus sequence variation in
RT tissues of rhesus macaques with simian AIDS."
RL J. Virol. 67:6522-6534(1993).
DR EMBL: L22792; AAA47643.1; -
DR INTERPRO: IPR000012; -
DR PFAM: PF00522; VPR; 1.
DR PRINTS: PR00444; HIVPRVXP.
SQ SEQUENCE 112 AA; 12846 MW; 7EC007A3BAFD7BEA CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;
Best Local Similarity 64.3%; Pred. No. 2.4;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 HWSYGLRGRGPPPP 15
| : | | | | | | |
Db 94 HGAGGWRGPPPP 107

RESULT 12
ID 088064 PRELIMINARY; PRT; 112 AA.
AC 088064;
DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 GN (MM3160) DNA, COMPLETE CDS'S.
 OS Chimpazee immunodeficiency virus (STV(cpz)) (CIV).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 RX NCBI_Taxid=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94016835; PubMed=8411355;
 RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;
 RT "Analysis of simian immunodeficiency virus sequence variation in
 RT tissues of rhesus macaques with simian AIDS.";
 RL J. Virol. 67:6522-6534(1993).
 DR EMBL: L22795; AAA47652.1; -;
 DR INTERPRO: IPR00012; -;
 DR PFAM: PF00522; VPR: 1.
 DR PRINTS: PR00444; HIVPRVPX.
 SQ SEQUENCE 112 AA; 12935 MW; D32070A4DAC1CF5 CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;
 Best Local Similarity 64.3%; Pred. No. 2.4;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HMSGRLRPGRP 15
 DB 94 HGAGGWRP 107

RESULT 13
 ID 088068 PRELIMINARY; PRT: 112 AA.
 AC 088068;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE (MM3160) DNA, COMPLETE CDS'S.
 GN VPX.
 OS Chimpazee immunodeficiency virus (STV(cpz)) (CIV).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 RX NCBI_Taxid=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94016835; PubMed=8411355;
 RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;
 RT "Analysis of simian immunodeficiency virus sequence variation in
 RT tissues of rhesus macaques with simian AIDS.";
 RL J. Virol. 67:6522-6534(1993).
 DR EMBL: L22797; AAA47658.1; -;
 DR INTERPRO: IPR00012; -;
 DR PFAM: PF00522; VPR: 1.
 DR PRINTS: PR00444; HIVPRVPX.
 SQ SEQUENCE 112 AA; 12841 MW; 8B82EF96CA137BF0 CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;
 Best Local Similarity 64.3%; Pred. No. 2.4;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HMSGRLRPGRP 15
 DB 94 HGAGGWRP 107

RESULT 14
 ID 088125 PRELIMINARY; PRT: 112 AA.
 AC 088125;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE VPX PROTEIN.

GN VPX.
 OS Chimpazee immunodeficiency virus (STV(cpz)) (CIV).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 RX NCBI_Taxid=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90272009; PubMed=1971917;
 RA Demhurst S., Embretson J.E., Anderson D.C., Mullins J.I., Fultz P.N.;
 RT "Sequence analysis and acute pathogenicity of molecularly cloned
 RT SIVSM-PBj14.";
 RL Nature 345:636-640(1990).
 DR EMBL: L03296; AAA47763.1; -;
 DR INTERPRO: IPR00012; -;
 DR PFAM: PF00522; VPR: 1.
 DR PRINTS: PR00444; HIVPRVPX.
 SQ SEQUENCE 112 AA; 12877 MW; BEADC7FF6CED2EAC CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;
 Best Local Similarity 64.3%; Pred. No. 2.4;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HMSGRLRPGRP 15
 DB 94 HGAGGWRP 107

RESULT 15
 ID 088131 PRELIMINARY; PRT: 112 AA.
 AC 088131;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE VPX PROTEIN.
 GN VPX.
 OS Chimpazee immunodeficiency virus (STV(cpz)) (CIV).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 RX NCBI_Taxid=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90272009; PubMed=1971917;
 RA Demhurst S., Embretson J.E., Anderson D.C., Mullins J.I., Fultz P.N.;
 RT "Sequence analysis and acute pathogenicity of molecularly cloned
 RT SIVSM-PBj14.";
 RL Nature 345:636-640(1990).
 DR EMBL: L03297; AAA47771.1; -;
 DR INTERPRO: IPR00012; -;
 DR PFAM: PF00522; VPR: 1.
 DR PRINTS: PR00444; HIVPRVPX.
 SQ SEQUENCE 112 AA; 12851 MW; 9C3EC7FF6CED3D14 CRC64;

Query Match 51.5%; Score 52; DB 12; Length 112;
 Best Local Similarity 64.3%; Pred. No. 2.4;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HMSGRLRPGRP 15
 DB 94 HGAGGWRP 107

Search completed: March 2, 2001, 11:00:49
 Job time: 459 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 11:01:20 ; Search time 32.03 Seconds
(without alignments)
16.132 Million cell updates/sec

Title: US-09-306-689-5
Perfect score: 101
Sequence: 1 XHMSYGLRPGRRPPPC 16

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues
Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	57.4	61	GON1_SHEEP	Q28588 ovine aries
2	58	57.4	63	GON1_MESAP	O09163 mesocricetu
3	58	57.4	67	GON1_MACMU	P55247 macaca mula
4	58	57.4	89	GON1_XENLA	P45656 xenopus lae
5	58	57.4	90	GON1_MOUSE	P13662 mus musculu
6	58	57.4	91	GON1_PIG	P49921 sus scrofa
7	58	57.4	92	GON1_HUMAN	P01148 homo sapien
8	58	57.4	92	GON1_RAT	P07490 rattus norv
9	58	57.4	92	GON1_TUPGB	O95335 tupiaa gils
10	54	53.5	10	GON1_ALAMI	P37041 alligator m
11	54	53.5	92	GON1_CHICK	P37042 gallus gall
12	52	51.5	94	GON1_HAPBU	P51918 haplochromi
13	52	51.5	95	GON1_PAGMA	P70074 pagrus majo
14	52	51.5	95	GON1_SPAUV	P51919 sparus majo
15	52	51.5	112	VPX_SIVM1	P05917 simian immu
16	52	51.5	112	VPX_SIVM2	P05918 simian immu
17	51	50.5	112	VPX_SIVM3	P05916 simian immu
18	51	50.5	112	VPX_SIVM4	P11266 simian immu
19	50	49.5	111	VPX_HV2KR	O74122 human immu
20	50	49.5	112	VPX_HV2CA	P24110 human immu
21	50	49.5	112	VPX_HV2D1	P17760 human immu
22	50	49.5	112	VPX_HV2RO	P06939 human immu
23	49	48.5	92	GON1_CAVPO	O54713 cavia porce
24	48	47.5	80	GON1_CLAGA	P33439 claritas gar
25	48	47.5	315	ALDR_PIG	P80276 sus scrofa
26	48	47.5	635	VP40_HSV11	P10210 herpes simp
27	47.5	47.0	114	GON2_TUPGB	O95336 tupiaa gils
28	47	46.5	436	SH6_RAT	P31388 rattus norv
29	47	46.5	2236	PVRL_DROME	P05990 drosophila
30	46.5	46.0	429	GAG_HTLIA	P03345 human t-cel
31	46.5	46.0	429	GAG_HTLIC	P14076 human t-cel
32	46.5	46.0	429	GAG_HTLIM	P14077 human t-cel
33	46.5	46.0	433	GAG_HTLV2	P03346 human t-cel

ALIGNMENTS

RESULT ID	1	GON1_SHEEP	STANDARD:	PRT:	61 AA.
AC	Q28588;				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	15-DEC-1998 (Rel. 37, Last annotation update)				
DE	PROGONADOLIBERIN I PRECURSOR (CONTAINS: GONADOLIBERIN I (LHRH I))				
DE	(LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING HORMONE I) (GNRH I) (LUTILIBERIN I); GNRH-ASSOCIATED PEPTIDE I)				
DE	(FRAGMENT).				
GN	GNRH1 OR GNRH OR LHRH.				
OS	Ovis aries (Sheep).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Caprinae; Ovis.				
RN	[1]				
RP	SEQUENCE OF 12-61 FROM N.A.				
RC	STRAIN-WESTERN RANGE; TISSUE-HYPOTHALAMUS;				
RA	Rodriguez R.E., Wise M.E.;				
RL	Submitted (Oct-1993) to the EMBL/Genbank/DBD databases.				
RN	[2]				
RP	SEQUENCE OF 1-10.				
RX	MEDLINE-72094314; PubMed-4550508;				
RA	Burgus R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,				
RA	Fellows R., Blackwell R., Vale W., Guillemin R.;				
RT	*Primary structure of the ovine hypothalamic luteinizing hormone-				
RT	releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass				
RT	spectrometry-decapeptide-Edman degradation).";				
RL	Proc. Natl. Acad. Sci. U.S.A. 69:278-282(1972).				
CC	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES				
CC	THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING				
CC	HORMONES.				
CC	-1- SIMILARITY: BELONGS TO THE GNRH FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL: U02517; AAA03433.1; -				
DR	PIR: A93780; RSHSG.				
DR	INTERPRO: IPR002012; -				
DR	PFAM: PF00446; GNRH; 1.				
DR	PROSITE: PS00473; GNRH; 1.				
KW	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;				
KW	Placenta.				
FT	NON_TER	1			
FT	CHAIN	1	>61		PROGONADOLIBERIN I.
FT	PEPTIDE	1	10		GONADOLIBERIN I.
FT	PEPTIDE	14	>61		GNRH-ASSOCIATED PEPTIDE I.
FT	ACT_SITE	3	3		APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT					ACTIVITY.

P50617 rattus norv
P20367 oncorhynchu
P55246 oncorhynchu
O92097 oncorhynchu
P30973 oncorhynchu
P35629 salmo salar
P45653 salmo trutt
P51922 porichthys
P45652 haplochromi
P51921 pagrus majo
P51923 sparus aura
P51917 carassius a

FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).
 FT NON_TER 61
 SO SEQUENCE 61 AA: 6828 MW: 63962A1AE319B8F0 CRC64;

Query Match 57.4%; Score 58; DB 1; Length 61;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 Db 2 HWSYGLRPG 10

RESULT 2

CON1_MESAU STANDARD; PRT; 63 AA.

AC 009163;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)]
 DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]
 DE (FRAGMENT).
 GN GNRH1 OR GNRH OR LHRH.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Mesocricetus.
 RN [1]
 RP Jansen H.T., Stevens P.J., Zeitler P., Lehman M.N.;
 RA Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; U91938; AAB51302.1; -;
 DR INTERPRO; IPR002012; -;
 DR PFAM; PF00446; GNRH; 1.
 DR PROSITE; PS00473; GNRH; 1.
 KM Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Placenta.
 FT NON_TER 1 1
 FT CHAIN 1 >63
 FT PEPTIDE 1 10
 FT PEPTIDE 14 >63
 FT ACT_SITE 3 3
 FT MOD_RES 1 1
 FT MOD_RES 10 10
 FT MOD_RES 63 63
 FT NON_TER 63 63
 SO SEQUENCE 63 AA: 7370 MW: FC94995676F77180 CRC64;

Query Match 57.4%; Score 58; DB 1; Length 63;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 Db 2 HWSYGLRPG 10

RESULT 3

CON1_MACMU STANDARD; PRT; 67 AA.

AC P55247;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)]
 DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]
 DE (FRAGMENT).
 GN GNRH1 OR GNRH OR LHRH.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-HYPOTHALAMUS;
 RX MEDLINE=95124501; Pubmed=7545971;
 RA Ma Y.J., Costa M.E., Ojeda S.R.;
 RT "Developmental expression of the genes encoding transforming growth factor alpha and its receptor in the hypothalamus of female rhesus macaques."
 RL Neuroendocrinology 60:346-359(1994).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; S75918; AAB33096.1; -;
 DR INTERPRO; IPR002012; -;
 DR PFAM; PF00446; GNRH; 1.
 DR PROSITE; PS00473; GNRH; 1.
 KM Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Signal.
 FT NON_TER 1 1
 FT SIGNAL <1 5
 FT CHAIN 6 >67
 FT PEPTIDE 6 15
 FT PEPTIDE 19 >67
 FT ACT_SITE 8 8
 FT MOD_RES 6 6
 FT MOD_RES 15 15
 FT MOD_RES 67 67
 FT NON_TER 67 67
 SO SEQUENCE 67 AA: 7573 MW: 505394DAA261A3F2 CRC64;

Query Match 57.4%; Score 58; DB 1; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 Db 7 HWSYGLRPG 15

```

RESULT 4
GONL_XENLA STANDARD: PRT: 89 AA.
AC P45656;
DR 01-NOV-1995 (Rel. 32, Created)
DR 01-NOV-1995 (Rel. 32, Last sequence update)
DR 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE (LH-RH) (LULIBERIN I).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-FOREBRAIN;
RA MEDLINE=94185563; PubMed=8137750;
RA Hayes W.P., Wray S., Battey J.F.;
RT "The frog gonadotropin-releasing hormone-I (GNRH-I) gene has a
RT mammalian-like expression pattern and conserved domains in
RT GNRH-associated peptide, but brain onset is delayed until
RT metamorphosis."
RL Endocrinology 134:1835-1844(1994).
CC -I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -I- SIMILARITY: BELONGS TO THE GNRH FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L28040; AAA9728.1;
DR INTERPRO; IPR002012;
DR PFWAM; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 23
FT CHAIN 24 89
FT PEPTIDE 24 33
FT CHAIN 37 89
FT PEPTIDE 37 89
FT MOD_RES 24 24
FT MOD_RES 33 33
FT MOD_RES 33 33
SQ SEQUENCE 89 AA; 10246 MW; 6FAF36FBAED04284 CRC64;

Query Match 57.4%; Score 58; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 HMSGYLRPG 10
DB 25 HMSGYLRPG 33

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87069928; PubMed=3024317;
RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
RA Phillips H.S., Nikolic K., Seeburg P.H.;
RT "A deletion truncating the gonadotropin-releasing hormone gene is
RT responsible for hypogonadism in the hpg mouse."
RL Science 234:1366-1371(1986).
CC -I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -I- SIMILARITY: BELONGS TO THE GNRH FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M14872; AAA37717.1;
DR MGD; MGI:95769; GNRH.
DR INTERPRO; IPR002012;
DR PFWAM; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 21
FT CHAIN 22 90
FT PEPTIDE 22 31
FT PEPTIDE 35 90
FT ACT_SITE 24 24
FT MOD_RES 22 22
FT MOD_RES 31 31
FT MOD_RES 31 31
SQ SEQUENCE 90 AA; 10337 MW; 1C0766FA4826E4D9 CRC64;

Query Match 57.4%; Score 58; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 HMSGYLRPG 10
DB 23 HMSGYLRPG 31

RESULT 6
GONL_PIG STANDARD: PRT: 91 AA.
AC P49921;
DR 01-OCT-1996 (Rel. 34, Created)
DR 01-OCT-1996 (Rel. 34, Last sequence update)
DR 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR (CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I).
CN GNRH1 OR GNRH.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-HYPOTHALAMUS;
RA Weesner G.D., Mattern R.L., Becker B.A.;
RL Submitted (May-1994) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 24-33.
RX MEDLINE=72114303; PubMed=4946067;
RA Baba Y., Matsuo H., Schnally A.V.;

```

```

RT  "Structure of the porcine LH- and FSH-releasing hormone. II.
RT  Confirmation of the proposed structure by conventional sequential
RT  analyses."
RL  Biochem. Biophys. Res. Commun. 44:459-463(1971).
RN  [3]
RP  SYNTHESIS OF GONADOLIBERIN.
RX  MEDLINE=72065376; PubMed=4942726;
RA  Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
RT  "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
RT  phase method."
RL  Biochem. Biophys. Res. Commun. 45:822-827(1971).
RN  [4]
RP  SYNTHESIS OF GONADOLIBERIN.
RX  MEDLINE=72117544; PubMed=4946275;
RA  Baba Y., Arimura A., Schally A.V.;
RT  "On the tryptophan residue in porcine LH and FSH-releasing hormone."
RL  Biochem. Biophys. Res. Commun. 45:483-487(1971).
CC  -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC  THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC  HORMONES.
CC  -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: L32864; AAA31066.1; -.
CC  DR  PIR: A01411; RHPGG.
CC  DR  INTERPRO: IPR002012; -.
CC  DR  PIR: PF00446; GNRH.1.
CC  DR  PROSITE: PS00473; GNRH.1.
CC  KW  Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC  Placenta; Signal.
CC  FT  CHAIN 1 23 PROGONADOLIBERIN I.
CC  FT  PEPTIDE 24 33 GONADOLIBERIN I.
CC  FT  ACT_SITE 34 91 GNRH-ASSOCIATED PEPTIDE I.
CC  FT  ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
CC  ACTIVITY.
CC  FT  MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
CC  FT  MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
CC  SQ  SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;

```

```

RP  SEQUENCE FROM N.A.
RX  MEDLINE=8936682; PubMed=2671939;
RA  Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT  "The complete nucleotide sequence of the human gonadotropin-releasing
RT  hormone gene."
RL  Nucleic Acids Res. 17:6403-6403(1989).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=86094338; PubMed=2867548;
RA  Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT  "Isolation of the gene and hypothalamic cDNA for the common precursor
RT  of gonadotropin-releasing hormone and prolactin release-inhibiting
RT  factor in human and rat."
RL  Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN  [3]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=85012739; PubMed=6090951;
RA  Seeburg P.H., Adelman J.P.;
RT  "Characterization of cDNA for precursor of human luteinizing hormone
RT  releasing hormone."
RL  Nature 311:666-668(1984).
RN  [4]
RP  SEQUENCE OF 24-33.
RX  MEDLINE=83126573; PubMed=6760865;
RA  Tan L., Rousseau P.;
RT  "The chemical identity of the immunoreactive LHRH-like peptide
RT  biosynthesized in the human placenta."
RL  Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
CC  -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC  THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC  HORMONES.
CC  -1- PHARMACEUTICAL: AVAILABLE UNDER THE NAMES FACTREL (AYERST LABS),
CC  LUTREPUSE OR LUTRELEF (FERRING PHARMACEUTICALS) AND RELISORM
CC  (SERONO).
CC  -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: X01059; CAA25526.1; -.
CC  DR  EMBL: M12578; AAA35916.1; -.
CC  DR  EMBL: X15215; CAA33285.1; -.
CC  DR  PIR: A01410; RHPGG.
CC  DR  PIR: A26173; A26173.
CC  DR  PIR: S05308; S05308.
CC  DR  MIR: 152760; -.
CC  DR  INTERPRO: IPR002012; -.
CC  DR  PIR: PF00446; GNRH.1.
CC  DR  PROSITE: PS00473; GNRH.1.
CC  KW  Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC  Placenta; Pharmaceutical; Signal.
CC  FT  CHAIN 1 23 PROGONADOLIBERIN I.
CC  FT  PEPTIDE 24 33 GONADOLIBERIN I.
CC  FT  ACT_SITE 34 91 GNRH-ASSOCIATED PEPTIDE I.
CC  FT  ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
CC  ACTIVITY.
CC  FT  MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
CC  FT  MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
CC  FT  CONFLICT 16 16 W->S (IN REF. 3).
CC  SQ  SEQUENCE 92 AA; 10380 MW; 30A7221B076FA79 CRC64;

```

```

Query Match          57.4%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33

```


Db 25 HWSYGLRPG 33

```

RESULT 8
ID GON1_RAT STANDARD: PRT; 92 AA.
AC P07490;
DT 01-APR-1988 (rel. 07, Last sequence update)
DT 01-APR-1988 (rel. 07, Last sequence update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR
DE I]
GN GNRH1 OR GNRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89384661; PubMed=2476669;
RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT "The rat gonadotropin-releasing hormone: SH locus: structure and
RT hypothalamic expression.";
RL Mol. Endocrinol. 3:1257-1262(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=THYMUS;
RX MEDLINE=93105480; PubMed=1468115;
RA Maier C.C., Marchetti B., Leboeuf R.D., Blalock J.E.;
RT "Thymocytes express a mRNA that is identical to hypothalamic
RT luteinizing hormone-releasing hormone mRNA.";
RL Cell. Mol. Neurobiol. 12:447-454(1992).
RN [4]
RP SEQUENCE OF 1-47 FROM N.A.
RC TISSUE=HEART;
RX MEDLINE=87149087; PubMed=3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
RT DNA locus.";
RL Science 235:1514-1517(1987).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S50870; AAB24572.1; -
CC EMBL: M12579; AAA41263.1; -
CC EMBL: M15670; AAA41264.1; -
CC EMBL: M15527; AAA42141.1; -
CC EMBL: M15529; AAA42139.1; -
CC EMBL: M15528; -; NOT_ANNOTATED_CDS.
CC PIR: B26173; RHRTG.
CC INTERPRO: IPR002012; -

```

```

DR PFAM: PF00446; GNRH: 1.
DR PROSITE: PS00473; GNRH: 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DABA3EB3 CRC64;
Query Match 57.4%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33
RESULT 9
ID GON1_TUPGB STANDARD: PRT; 92 AA.
AC Q95335;
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 30-MAY-2000 (rel. 39, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
GN GNRH1 OR GNRH.
OS Tupia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupala.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HYPOTHALAMUS;
RX MEDLINE=97079639; PubMed=8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGNRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GNRH gene expression in a
RT placental mammal.";
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U63326; AAB16837.1; -
CC INTERPRO: IPR002012; -
CC PFAM: PF00446; GNRH: 1.
DR PROSITE: PS00473; GNRH: 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
BY SIMILARITY.
PROGNADOLIBERIN I.
GONADOLIBERIN I.
GNRH-ASSOCIATED PEPTIDE I.
APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
ACTIVITY.

```

```

FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
FT STRAIN-WHITE LEGHORN;
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
FT SEQUENCE 92 AA; 10197 MW; 4FDB2C58CF5F63B CRC64;

Query Match
Best Local Similarity 57.4%; Score 58; DB 1; Length 92;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33

RESULT 10
GONI_ALLMT STANDARD; PRT; 10 AA.
ID GONI_ALLMT
AC P37041; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GONADOLIBERIN I (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I) (LH-RH I)
DE (LULIBERIN I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorineae; Alligator.
RN [1]
RP SEQUENCE.
RC TISSUE-BRAIN;
RX MEDLINE=91352338; PubMed=1882082;
RA Lovejoy D.A., Fischer W.H., Parker D.B., Moroy J.E., Park M.,
  Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
  "Primary structure of two forms of gonadotropin-releasing hormone
  from brains of the American alligator (Alligator mississippiensis).";
  Regul. Pept. 33:105-116(1991).
CC CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR PIR: A60066; RNRH.1.
DR INTERPRO: IPR002012; -.
DR PIR: PFO0446; GNRH. 1.
DR PROSITE: PS00473; GNRH. 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;

Query Match
Best Local Similarity 53.5%; Score 54; DB 1; Length 10;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 10

RESULT 11
GONI_CHICK STANDARD; PRT; 92 AA.
ID GONI_CHICK
AC P37042; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
  DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
  HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
  Gallus.
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEGHORN;
RX MEDLINE=94059355; PubMed=7902095;
RA Dunn I.C., Chen Y., Hook C., Sharp P.J., Sang H.M.;
  "Characterization of the chicken preprogonadotropin-releasing
  hormone-1 gene.";
  J. Mol. Endocrinol. 11:19-29(1993).
RN [2]
RP SEQUENCE OF 24-33.
RC TISSUE-HYPOTHALAMUS;
RX MEDLINE=82265778; PubMed=7050119;
RA King J.A., Millar R.P.;
  "Structure of chicken hypothalamic luteinizing hormone-releasing
  hormone. II. Isolation and characterization.";
  J. Biol. Chem. 257:10729-10732(1982).
RN [3]
RP SEQUENCE OF 24-33.
RC TISSUE-HYPOTHALAMUS;
RA King J.A., Millar R.P.;
  "Structure of avian hypothalamic gonadotropin-releasing hormone.";
  S. Afr. J. Sci. 78:124-125(1982).
RN [4]
RP SYNTHESIS OF 24-33.
RX MEDLINE=82265777; PubMed=7050118;
RA King J.A., Millar R.P.;
  "Structure of chicken hypothalamic luteinizing hormone-releasing
  hormone. I. Structural determination on partially purified
  material.";
  J. Biol. Chem. 257:10722-10728(1982).
CC CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
  CC or send an email to license@isb-sib.ch).
  CC -----
DR EMBL: X69491; CAA49246.1; -.
DR PIR: S33507; S33507.
DR INTERPRO: IPR002012; -.
DR PIR: PFO0446; GNRH. 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
  Signal.
FT SIGNAL 1 23
FT CHAIN 24 92 PROGONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE 92 AA; 10206 MW; 61AEB7EBAF508B6A CRC64;

Query Match
Best Local Similarity 53.5%; Score 54; DB 1; Length 92;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33

RESULT 12
GONI_HAPBU STANDARD; PRT; 94 AA.
ID GONI_HAPBU
AC P51918; O93387;
DT 01-OCT-1996 (Rel. 34, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)

```

DE (LH-RH I) (LULIBERIN I).
GN GnRH.
OS Haplochromis burtoni.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;
OC Cichlidae; Astatotilapia.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95396797; PubMed-7667296;
RA White S.A., Kastan T.L., Bond C.T., Adelman J.P., Fernald R.D.;
RT "Three gonadotropin-releasing hormone genes in one organism suggest
RT novel roles for an ancient peptide."
RL Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99061842; PubMed-9843638;
RA White R.B., Fernald R.D.;
RT "Ontogeny of gonadotropin-releasing hormone (GnRH) gene expression
RT reveals a distinct origin for GnRH-containing neurons in the
RT midbrain."
RL Gen. Comp. Endocrinol. 112:322-329(1998).
RN [3]
RP SEQUENCE OF 23-32.
RC TISSUE-PITUITARY;
RX MEDLINE-95372591; PubMed-7644702;
RA Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,
RA White S.A., Francis R.C., Fernald R.D., Licht P., Wardy C.,
RA Sherwood N.M.;
RT "Primary structure of solitary form of gonadotropin-releasing hormone
RT (GnRH) in cichlid pituitary; three forms of GnRH in brain of cichlid
RT and pumpkinseed fish."
RL Regul. Pept. 57:43-53(1995).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE
CC RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALAMIC-PITUITARY-
CC GONADAL AXIS.
CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS
CC TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPOTHYSEAL AXONS.
CC -1- MASS SPECTROMETRY: MW-1113.9; METHOD-MALDI; RANGE-23-32.
CC -1- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U31865; AAC59691.1; -;
DR EMBL: AF076961; AAC27716.1; -;
DR INTERPRO: IPR002012; -;
DR PFM: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
RN Signal; Multigene family.
FT SIGNAL 1 22
FT CHAIN 23 94
FT PEPTIDE 23 94
FT MOD_RES 36 94
FT MOD_RES 23 23
FT MOD_RES 32 32
FT MOD_RES 86 94
FT MOD_RES 10382 MW; E57DBA8333278D7 CRC64;
FT SEQUENCE 94 AA; 10382 MW; E57DBA8333278D7 CRC64;
Query Match 51.5%; Score 52; DB 1; Length 94;
Best Local Similarity 88.9%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 13
GONL_PAGMA STANDARD; PRT; 95 AA.
ID GONL_PAGMA
AC P70074;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE (LH-RH I) (LULIBERIN I).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
OC Sparidae; Chrysophrys.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX Okuzawa K., Graneman J., Bogerd J., Goos H., Zohar Y., Kagawa H.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GnRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D86582; BAA13129.1; -;
DR INTERPRO: IPR002012; -;
DR PFM: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
RN Signal; Multigene family.
FT SIGNAL 1 23
FT CHAIN 24 95
FT PEPTIDE 24 95
FT PEPTIDE 37 95
FT MOD_RES 24 24
FT MOD_RES 33 33
FT MOD_RES 61E79C990328D73E CRC64;
FT SEQUENCE 95 AA; 10566 MW; 61E79C990328D73E CRC64;
Query Match 51.5%; Score 52; DB 1; Length 95;
Best Local Similarity 88.9%; Pred. No. 1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 14
GONL_SPRAU STANDARD; PRT; 95 AA.
ID GONL_SPRAU
AC P51919;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE (LH-RH I) (LULIBERIN I) (SBCNRH).
OS Sparus aurata (Gilthead sea bream).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
OC Sparidae; Sparus.

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN:
RX MEDLINE=95268499; PubMed=7749463;
RA Gotthelf Y., Elizur A., Chow M., Chen T.T., Zohar Y.;
RT "Molecular cloning and characterization of a novel gonadotropin-
  releasing hormone from the gillhead seabream (Sparus aurata).";
  Mol. Mar. Biol. Biotechnol. 4:27-35(1995).
RN [2]
RP SEQUENCE OF 26-35.
RC TISSUE-BRAIN:
RX MEDLINE=95083645; PubMed=7991588;
RA Powell J.F.F., Zohar Y., Elizur A., Park M., Fischer W.H.,
  Craig A.G., Rivier J.E., Lovejoy D.A., Sherwood N.M.;
RT "Three forms of gonadotropin-releasing hormone characterized from
  brains of one species.";
  Proc. Natl. Acad. Sci. U.S.A. 91:12081-12085(1994).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- MASS SPECTROMETRY: MW=1113.6; METHOD=MALDI; RANGE=26-35.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U30320; AAA75469.1; -.
DR INTERPRO: IPR002012; -.
DR PFM: PF00446; GNRH. 1.
DR PROSITE: PS00473; GNRH. 1.
KM Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
  Signal; Multigene family.
FT SIGNAL 1 25
FT CHAIN 26 95 PROGONADOLIBERIN I.
FT PEPTIDE 26 95 GONADOLIBERIN I.
FT MOD_RES 39 95 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
FT MOD_RES 26 26 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 35 35 AMIDATION (G-36 PROVIDE AMIDE GROUP).
SQ SEQUENCE 95 AA; 10753 MW; 49313FD6FDB87DA CRC64;

Query Match
Best Local Similarity 51.5%; Score 52; DB 1; Length 95;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
   |||||
Db 27 HWSYGLSPG 35

RESULT 15
VPX_SIVM1 STANDARD; PRT; 112 AA.
AC P05917;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-OCT-1989 (Rel. 12, Last annotation update)
DE VPX PROTEIN (X ORF PROTEIN).
GN VPX.
OS Simian immunodeficiency virus (Mm142-83 isolate) (SIV-MAC).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87287230; PubMed=3649576;
RA Chakrabarti L., Guyader M., Allison M., Daniel M.D., Desrosiers R.C.,
  Tiollais P., Sonigo P.;
RT "Sequence of simian immunodeficiency virus from macaque and its
  relationship to other human and simian retroviruses.";
  Nature 328:543-547(1987).
CC -1- MISCELLANEOUS: THIS IS A MACAQUE ISOLATE.
```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y00277; CAA68382.1; -.
DR PIR: D28887; ASLJX3.
DR HIV: M16403; VPXMM142.
DR INTERPRO: IPR000012; -.
DR PFM: PF00522; VPR. 1.
DR PRINTS: PR00444; HIVVPRVPX.
KW AIDS.
SQ SEQUENCE 112 AA; 12906 MW; 46F564F45AFED960 CRC64;

Query Match
Best Local Similarity 51.5%; Score 52; DB 1; Length 112;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HWSYGLRPPRPP 15
   | : | | | | | |
Db 94 HGAGGWRPGRPP 107

Search completed: March 2, 2001, 11:01:20
Job time: 435 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:55:40 ; Search time 56.87 Seconds
(without alignments)
19.103 Million cell updates/sec

Title: US-09-306-689-5

Perfect score: 101

Sequence: 1 XHMSYGLRPGRRPPPC 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	58	57.4	10	1	RHPGC	gonadoliberin - pi
2	58	57.4	10	1	RHSHG	gonadoliberin - sh
3	58	57.4	67	2	I78541	gonadoliberin precursor
4	58	57.4	89	2	I51423	gonadoliberin precursor
5	58	57.4	90	1	RHMSG	gonadoliberin precursor
6	58	57.4	92	1	RHHUG	gonadoliberin precursor
7	58	57.4	92	1	RHRTG	gonadoliberin precursor
8	55	54.5	969	2	T15446	gonadoliberin precursor
9	54	53.5	10	1	RHAQ1	gonadoliberin I -
10	54	53.5	92	2	I50644	gonadoliberin I precursor
11	52	51.5	98	2	I50739	gonadotropin-releasing
12	52	51.5	106	2	A30996	gonadotropin-releasing
13	52	51.5	112	1	ASLJX3	gonadotropin-releasing
14	52	51.5	112	2	S03066	gonadotropin-releasing
15	51.5	112	1	ASLJST	gonadotropin-releasing	
16	50	49.5	112	1	ASLJX2	gonadotropin-releasing
17	50	49.5	112	1	ASLJX2	gonadotropin-releasing
18	50	49.5	112	2	S12155	gonadotropin-releasing
19	50	49.5	112	2	S53094	gonadotropin-releasing
20	50	49.5	112	2	T32976	gonadotropin-releasing
21	50	49.5	112	2	A56694	gonadotropin-releasing
22	49	48.5	373	2	A47234	gonadotropin-releasing
23	49	48.5	2129	2	T27431	gonadotropin-releasing
24	48	47.5	80	1	RHIDS	gonadoliberin I precursor
25	48	47.5	316	2	A59021	gonadoliberin I precursor
26	48	47.5	598	2	T48822	gonadoliberin I precursor
27	48	47.5	629	2	S31174	gonadoliberin I precursor
28	47	46.5	635	1	NMBEW6	gonadoliberin I precursor
29	47	46.5	436	2	JN0591	gonadoliberin I precursor

30	47	46.5	587	2	T19893 hypothetical prote
31	47	46.5	587	1	T29324 hypothetical prote
32	47	46.5	2236	1	QZFF rudimentary protei
33	46.5	46.0	429	1	FOLJGH probable gag poly
34	46.5	46.0	429	1	FOLJCN gag polypoteins -
35	46.5	46.0	429	1	S06073 gag polypoteins -
36	46.5	46.0	433	1	FOLJH2 gag polypoteins -
37	46	45.5	477	2	F64805 RhsA protein (lipo
38	46	45.5	972	2	T47922 anthranilate phosph
39	45.5	45.0	112	2	S19976 Ig kappa chain V r
40	45.5	45.0	613	2	T42671 hypothetical prote
41	45	44.6	10	2	A21114 gonadotropin-relea
42	45	44.6	74	2	I51092 gonadotropin-relea
43	45	44.6	82	2	I51180 gonadotropin-relea
44	45	44.6	82	2	I51355 gonadotropin-relea
45	45	44.6	82	2	I51365 gonadotropin-relea

ALIGNMENTS

```
RESULT 1
RHPGC
gonadoliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411
R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of
A:Reference number: A90172; MUID:72114303
A:Accession: A01411
A:Molecule type: protein
A:Residues: 1-10 <BAB>
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me
A:Reference number: A90176; MUID:72065376
A:Contents: annotation; synthesis
A>Note: The synthetic and natural hormones have the same physicochemical and biologic
R:Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A:Reference number: A90175; MUID:72117544
A:Contents: annotation
A>Note: Trp-3 appears to be essential for biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
F10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 57.4%; Score 58; DB 1; Length 10;
Best local similarity 100.0%; Pred. No. 0.038;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HMSYGLRPG 10
DB 2 HMSYGLRPG 10

RESULT 2
RHSHG
gonadoliberin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C:Accession: A93780; A01411
R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa
A:Reference number: A93780; MUID:72094314
A:Accession: A93780
```

A:Molecule type: protein
A:Residues: 1-10 <8UR>
A>Note: the natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormone
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 57.4%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
|||||
DB 2 HWSYGLRPG 10

RESULT 3
178541
gonadoliberin precursor - rhesus macaque (fragment)
N:Alternate names: luteinizing hormone releasing hormone
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C:Accession: I78541
R:Ma, Y.T.; Costa, M.E.; Ojeda, S.R.
Neuroendocrinology 60, 346-359, 1994
A:Title: Developmental expression of the genes encoding transforming growth factor alpha
A:Reference number: I58134; MUID:95124501
A:Accession: I78541
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-67 <RES>
A:Cross-references: GB:S75918; NID:g912831; PIDN:AAA33096.1; PID:g912832
C:Superfamily: gonadoliberin

Query Match 57.4%; Score 58; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
|||||
DB 7 HWSYGLRPG 15

RESULT 4
151423
gonadoliberin precursor - African clawed frog
N:Alternate names: luteinizing hormone releasing hormone
C:Species: Xenopus laevis (African clawed frog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: I51423
R:Hayes, W.P.; Wray, S.; Battey, J.F.
Endocrinology 134, 1835-1845, 1994
A:Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved domain
A:Reference number: I51423; MUID:94185563
A:Accession: I51423
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-89 <HAY>
A:Cross-references: GB:L28040; NID:g496291; PIDN:AAA9728.1; PID:g496292
C:Genetics:
A:Gene: GnRH-I
C:Superfamily: gonadoliberin

Query Match 57.4%; Score 58; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10

DB 25 HWSYGLRPG 33
|||||

RESULT 5
RHMSG
gonadoliberin precursor - mouse
N:Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing hormone (LHRH)
C:Species: Mus musculus (house mouse)
C>Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
C:Accession: A47578
R:Wason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikoli
Science 234, 1366-1371, 1986
A:Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible for the hypogonadism in the mouse
A:Reference number: A47578; MUID:87069928
A:Accession: A47578
A:Molecule type: DNA
A:Residues: 1-90 <MAS>
A:Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387175
C:Genetics:
A:Introns: 45/3; 77/3
C:Function:
A:Description: gonadoliberin stimulates pituitary secretion of luteinizing hormone and follicle stimulating hormone
A>Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:22-31/Product: gonadoliberin #status predicted <GLR>
F:35-90/Product: gonadoliberin-associated protein #status predicted <GAP>
F:22/Modified site: pyroglutamate carboxylic acid (Gln) (in mature form) #status predicted
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 57.4%; Score 58; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
|||||
DB 23 HWSYGLRPG 31

RESULT 6
RHMSG
gonadoliberin precursor - human
N:Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasing hormone (LHRH)
C:Species: Homo sapiens (man)
C>Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 18-Jul-1999
C:Accession: S05308; A26173; A93442; A90108; A01410; S45718
R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
Nucleic Acids Res. 17, 6403-6404, 1989
A:Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone gene
A:Reference number: S05308; MUID:89366682
A:Accession: S05308
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-92 <HAY>
A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and luteinizing hormone-releasing hormone
A:Reference number: A94090; MUID:86094338
A:Accession: A26173
A:Molecule type: mRNA
A:Residues: 1-92 <ADE>
A:Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749
R:Seeburg, P.H.; Adelman, J.P.
Nature 311, 666-668, 1984
A:Title: Characterization of cDNA for precursor of human luteinizing hormone releasing hormone
A:Reference number: A93342; MUID:85012739

A:Accession: A93342
A:Molecule type: mRNA
A:Residues: 1-15, 'S', 17-92 <SEE>
A:Cross-references: GB:X01059; NID:934356; PIDN:CAA25526.1; PID:934357
A:Experimental source: placenta
R:Tan, L.; Rousseau, P.
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982
A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in
A:Reference number: A90108; MUID:83126573
A:Accession: A90108
A:Molecule type: protein
A:Residues: 24-33 <TAN>
A:Experimental source: placental trophoblasts
R:Leibovitz, D.; Koch, Y.; Piltzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amsterda
FBS Lett. 346, 203-206, 1994
A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by th
A:Reference number: S45718; MUID:94283597
A:Contents: annotation; degradation pathway of synthetic hormone
C:Genetics:
A:Gene: GDB:GNRH; LHRH; GRH
A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760
A:Map position: 8p21-8p11.2
A:Introns: 47/3; 79/3
C:Function:
A:Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitropin
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyrogutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-92/Product: progadoliberin #status predicted <PGN>
F:24-33/Product: gonadoliberin #status experimental <MAT>
F:37-92/Product: gonadoliberin-associated protein #status predicted <GAP>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 57.4%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HMSGYLRPG 10
DB 25 HMSGYLRPG 33

RESULT 7
RHRGTG
gonadoliberin precursor - rat
N:Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hormo
N:Contains: gonadoliberin; prolactin release-inhibiting factor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
C:Accession: A40147; B26173; A48410
R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
MOL. Endocrinol. 3, 1257-1262, 1989
A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex
A:Reference number: A40147; MUID:89384661
A:Accession: A40147
A:Molecule type: DNA
A:Residues: 1-92 <BON>
A:Cross-references: GB:M31670; NID:9204447; PIDN:AAA1264.1; PID:9204448
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot
A:Reference number: A94090; MUID:86094338
A:Accession: B26173
A:Molecule type: mRNA
A:Residues: 1-92 <NDE>
A:Cross-references: GB:M12579; NID:9204445; PIDN:AAA1263.1; PID:9204446
R:Walter, C.C.; Marchetti, B.; LeBoeuf, R.D.; Bialock, J.E.
Cell. Mol. Neurobiol. 12, 447-454, 1992
A:Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone
A:Reference number: A48410; MUID:93105480

A:Accession: A48410
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <MAT>
A:Cross-references: GB:S50870; NID:9262059; PIDN:AAB24572.1; PID:9262060
A:Experimental source: thymus
A:Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBI:121083)
C:Genetics:
A:Introns: 47/3; 79/3
C:Function:
A:Description: stimulates pituitary secretion of luteotropin and follitropin
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activi
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyrogutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-92/Product: progadoliberin #status predicted <PGN>
F:24-33/Product: gonadoliberin #status predicted <GNT>
F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 57.4%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HMSGYLRPG 10
DB 25 HMSGYLRPG 33

RESULT 8
T15446
hypothetical protein C07G1.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15446
R:Hawkins, J.
submitted to the EMBL Data Library, May 1996
A:Description: The sequence of C. elegans cosmid C07G1.
A:Reference number: Z18352
A:Accession: T15446
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-969 <HAW>
A:Cross-references: EMBL:U58751; NID:91326379; PID:91326381; PIDN:AAB00657.1; GSPDB:G
A:Experimental source: strain Bristol N2; clone C07G1
C:Genetics:
A:Gene: CESP:C07G1.4
A:Map position: 4
A:Introns: 110/3; 149/2; 515/1; 564/1; 600/2; 690/2; 755/3; 843/3; 884/3; 940/2

Query Match 54.5%; Score 55; DB 2; Length 969;
Best Local Similarity 75.0%; Pred. No. 9.2;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 SYGLRPPRPP 15
DB 727 SHGLAPRPP 738

RESULT 9
RHAQI
gonadoliberin I - American alligator
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: A60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swan
Regul. Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains
A:Reference number: A60066; MUID:91352338

A:Accession: A60066
 A:Molecule type: protein
 A:Residues: 1-10 <LOV>
 C:Superfamily: gonadoliberin
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 53.5%; Score 54; DB 1; Length 10;
 Best Local Similarity 88.9%; Pred. No. 0.13;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 | | | | | | | |
 Db 2 HWSYGLQPG 10

RESULT 10
 I50644

gonadoliberin I precursor - chicken
 N:Alternate names: gonadotropin-releasing hormone I
 C:Species: Gallus gallus (chicken)
 C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
 C:Accession: I50644; S33507
 R:Dunn, I.C.; Chen, Y.; Hook, C.; Sharp, P.J.; Sang, H.M.
 J. Mol. Endocrinol. 11, 19-29, 1993
 A>Title: Characterization of the chicken preprogonadotropin-releasing hormone-I gene.
 A:Reference number: I50644; PMID:94059355
 A:Accession: I50644
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-92 <DD2>
 A:Cross-references: EMBL:X69491; NID:9496326; PIDN:CAA49246.1; PID:g311612
 C:Genetics: 47/3; 79/3
 A:Introns: 47/3; 79/3
 C:Superfamily: gonadoliberin

Query Match 53.5%; Score 54; DB 2; Length 92;
 Best Local Similarity 88.9%; Pred. No. 1.2;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 | | | | | | | |
 Db 25 HWSYGLQPG 33

RESULT 11

I50739
 gonadotropin-releasing hormone - Cichlid (Haplochromis burtoni)
 C:Species: Haplochromis burtoni
 C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C:Accession: I50739
 R:White, S.A.; Kasten, T.L.; Bond, C.T.; Adelman, J.P.; Fernald, R.D.
 Proc. Natl. Acad. Sci. U.S.A. 92, 8363-8367, 1995
 A>Title: Three gonadotropin-releasing hormone genes in one organism suggest novel roles
 A:Reference number: I50739; PMID:95396797
 A:Accession: I50739
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-98 <WHI>
 A:Cross-references: EMBL:U11865; NID:9905398; PIDN:AAC59691.1; PID:g905399
 C:Superfamily: gonadoliberin

Query Match 51.5%; Score 52; DB 2; Length 98;
 Best Local Similarity 88.9%; Pred. No. 2.3;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 | | | | | | | |
 Db 24 HWSYGLSPG 32

RESULT 12

A30996
 orf-X protein - simian immunodeficiency virus (isolate Mne) (fragments)
 N:Alternate names: vpx protein
 C:Species: simian immunodeficiency virus, SIV
 C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Apr-1994
 C:Accession: A30996
 R:Henderson, L.E.; Sowder, R.C.; Copeland, T.D.; Benveniste, R.E.; Oroszian, S.
 Science 241, 199-201, 1988
 A>Title: Isolation and characterization of a novel protein (X-ORF product) from SIV a
 A:Reference number: A30996; PMID:88264422
 A:Accession: A30996
 A:Molecule type: protein
 A:Residues: 1-106 <HEN>
 C:Genetics:
 A:Gene: vpx
 C:Superfamily: AIDS vpu protein

Query Match 51.5%; Score 52; DB 2; Length 106;
 Best Local Similarity 64.3%; Pred. No. 2.5;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HWSYGLRGRPPP 15
 | : | | | | | |
 Db 88 HGAGGWRGPPPP 101

RESULT 13

ASLJX3
 vpu protein - simian immunodeficiency virus (macaque isolate)
 N:Alternate names: orf-X protein
 C:Species: simian immunodeficiency virus, SIV
 C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Jul-1999
 C:Accession: D28887
 R:Chakrabarti, L.; Guyader, M.; Alizon, M.; Daniel, M.D.; Desrosiers, R.C.; Tiollais, P.
 Nature 328, 543-547, 1987
 A>Title: Sequence of simian immunodeficiency virus from macaque and its relationship
 A:Reference number: A28887; PMID:87287230
 A:Accession: D28887
 A:Molecule type: DNA
 A:Residues: 1-112 <CHA>
 A:Cross-references: GB:Y00277; GB:M16403; NID:961730; PIDN:CAA68382.1; PID:g61734
 C:Genetics:
 A:Gene: vpu
 C:Superfamily: AIDS vpu protein
 C:Keywords: AIDS; Immunodeficiency

Query Match 51.5%; Score 52; DB 1; Length 112;
 Best Local Similarity 64.3%; Pred. No. 2.6;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HWSYGLRGRPPP 15
 | : | | | | | |
 Db 94 HGAGGWRGPPPP 107

RESULT 14

S03066
 gene X protein - human T-cell lymphotropic virus type 4
 C:Species: human T-cell lymphotropic virus type 4, HTLV-4
 C>Date: 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
 C:Accession: S03066
 R:Hahn, B.H.; Kong, L.I.; Lee, S.W.; Kumar, P.; Taylor, M.E.; Arya, S.K.; Shaw, G.M.
 Nature 300, 184-186, 1987
 A>Title: Relation of HTLV-4 to simian and human immunodeficiency-associated viruses.
 A:Reference number: S03066
 A:Accession: S03066
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA

A:Residues: 1-112 <HAH>
A:Cross-references: EMBL:X06391; NID:g61580; PIDN:CAA29688.1; PID:g61581
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1987
C:Superfamily: AIDS vpu protein

Query Match 51.5%; Score 52; DB 2; Length 112;
Best Local Similarity 64.3%; Pred. No. 2.6;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 2 HWSYGLRPRGPPPP 15
I : | | | | |
Db 94 HGAGGWRGPPPP 107

RESULT 15
ASLJST
vpu protein - simian immunodeficiency virus SIVagm (type 3, isolate STRV-3agm)
N:Alternate names: orf-X protein
C:Species: simian immunodeficiency virus SIVagm
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 31-Jan-1997
C:Accession: D26737
R:Hirsch, V.
submitted to GenBank, June 1987
A:Reference number: A26739
A:Accession: D26737
A:Molecule type: DNA
A:Residues: 1-112 <HIR>
A:Genetics:
A:Gene: vpu
C:Superfamily: AIDS vpu protein
C:Keywords: AIDS; immunodeficiency

Query Match 50.5%; Score 51; DB 1; Length 112;
Best Local Similarity 64.3%; Pred. No. 3.5;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HWSYGLRPRGPPPP 15
I : | | | | |
Db 94 HRAGGWRGPPPP 107

Search completed: March 2, 2001, 10:55:41
Job time: 261 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:54:36 ; Search time 47.48 Seconds
(without alignments)
6.429 Million cell updates/sec

Title: US-09-306-689-4

Sequence: 1 XHMSYGLRPPSSPPPC 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents.AA:*
2: /cgn2_6/prodata/2/1aa/3A_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/PCPTUS_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/Packfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103	99.0	17	1	US-08-188-223-6
2	103	99.0	17	1	US-08-968-466-6
3	83.5	80.3	16	1	US-08-188-223-2
4	83.5	80.3	16	3	US-08-968-466-2
5	65	62.5	52	3	US-08-458-814-6
6	65	62.5	55	3	US-08-458-814-7
7	62	59.6	14	1	US-08-453-588-23
8	62	59.6	14	3	US-08-521-079-23
9	62	59.6	17	1	US-07-690-983D-18
10	62	59.6	18	1	US-07-690-983D-18
11	62	59.6	49	1	US-08-387-156-4
12	62	59.6	49	2	US-08-694-865-4
13	62	59.6	49	3	US-08-878-748-4
14	62	59.6	49	3	US-09-124-491-4
15	62	59.6	544	2	US-08-387-156-10
16	62	59.6	544	2	US-08-694-865-10
17	62	59.6	544	2	US-08-878-748-10
18	62	59.6	544	2	US-09-124-491-10
19	62	59.6	544	2	US-08-694-865-16
20	62	59.6	699	3	US-08-694-865-16
21	62	59.6	977	1	US-09-124-491-16
22	62	59.6	977	1	US-08-387-156-8
23	62	59.6	977	2	US-08-694-865-8
24	62	59.6	977	2	US-08-878-748-8
25	59	56.7	12	3	US-09-124-491-8
26	59	56.7	12	1	US-08-453-588-24
27	59	56.7	14	1	US-08-521-079-24
28	59	56.7	14	1	US-07-690-983D-22

29	59	56.7	14	1	US-07-690-983D-26	Sequence 26, Appl
30	59	56.7	24	1	US-07-690-983D-43	Sequence 43, Appl
31	59	56.7	44	1	US-07-690-983D-45	Sequence 45, Appl
32	59	56.7	84	1	US-07-690-983D-47	Sequence 47, Appl
33	58	55.8	10	1	US-07-714-540-9	Sequence 9, Appl
34	58	55.8	10	1	US-07-690-983D-2	Sequence 2, Appl
35	58	55.8	10	1	US-07-690-983D-32	Sequence 32, Appl
36	58	55.8	10	1	US-08-103-022-1	Sequence 1, Appl
37	58	55.8	10	1	US-08-184-935-6	Sequence 6, Appl
38	58	55.8	10	1	US-08-343-883-1	Sequence 1, Appl
39	58	55.8	10	1	US-08-000-931-5	Sequence 5, Appl
40	58	55.8	10	1	US-08-428-488-22	Sequence 11, Appl
41	58	55.8	10	1	US-08-341-219-11	Sequence 22, Appl
42	58	55.8	10	1	US-08-453-588-2	Sequence 2, Appl
43	58	55.8	10	1	US-08-453-588-4	Sequence 4, Appl
44	58	55.8	10	1	US-08-453-588-6	Sequence 6, Appl
45	58	55.8	10	5	5492893-1	Patent No. 5492893

ALIGNMENTS

RESULT 1
US-08-188-223-6
Sequence 6, Application US/08188223
Patent No. 5688506
GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
APPLICANT: Scibienski, Robert
TITLE OF INVENTION: Immunogens Against Gonadotropin
TITLE OF INVENTION: Releasing Hormone
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dimitrios T. Drivas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188,223
FILING DATE: 27-JAN-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8286
TELEFAX: 212-354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label=pcgu
FEATURE:
NAME/KEY: Region
LOCATION: 1..10
OTHER INFORMATION: /note="pyroglutamic acid (5-oxoproline)"
FEATURE:
NAME/KEY: Region

LOCATION: 11..17
OTHER INFORMATION: /note= "spacer"
US-08-188-223-6

Query Match 99.0%; Score 103; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPC 17
Db 2 HWSYGLRPGSSPPPC 17

RESULT 2
US-08-968-466-6
Sequence 6, Application US/08968466
Patent No. 6132720

GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
APPLICANT: Scibienski, Robert
TITLE OF INVENTION: Immunogens Against Gonadotropin
TITLE OF INVENTION: Releasing Hormone
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dimitrios T. Drivas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10036-2787

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968.466
FILING DATE: 27-JAN-1994
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8286
TELEFAX: 212-354-8113

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES

FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= pglu
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"

FEATURE:
NAME/KEY: Region
LOCATION: 1..10
OTHER INFORMATION: /note= "immunomimic"
FEATURE:
NAME/KEY: Region
LOCATION: 11..17
OTHER INFORMATION: /note= "spacer"
US-08-968-466-6

Query Match 99.0%; Score 103; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPC 17
Db 2 HWSYGLRPGSSPPPC 17

RESULT 3
US-08-188-223-2
Sequence 2, Application US/08188223
Patent No. 5688506

GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
APPLICANT: Scibienski, Robert
TITLE OF INVENTION: Immunogens Against Gonadotropin
TITLE OF INVENTION: Releasing Hormone
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dimitrios T. Drivas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY

COUNTRY: USA
ZIP: 10036-2787

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188.223
FILING DATE: 27-JAN-1994
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8286
TELEFAX: 212-354-8113

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES

FEATURE:
NAME/KEY: Region
LOCATION: 1..10
OTHER INFORMATION: /note= "immunomimic"
FEATURE:
NAME/KEY: Region
LOCATION: 11..16
OTHER INFORMATION: /note= "spacer"

FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= pglu
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"
US-08-188-223-2

Query Match 80.3%; Score 83.5; DB 1; Length 16;
Best Local Similarity 87.5%; Pred. No. 5.3e-05;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 2 HWSYGLRPGSSPPPC 17
Db 2 HWSYGLRPG-RRPPC 16

RESULT 4
US-08-968-466-2

```
Sequence 2, Application US/08968466
Patent No. 6132720
GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
APPLICANT: Scibienski, Robert
TITLE OF INVENTION: Immunogens Against Gonadotropin
TITLE OF INVENTION: Releasing Hormone
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dimitrios T. Divas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,466
FILING DATE: 27-JAN-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Divas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-354-8113
TELEFAX: 212-354-8113
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
FEATURE:
NAME/KEY: Region
LOCATION: 1..10
OTHER INFORMATION: /note= "immunomimic"
FEATURE:
NAME/KEY: Region
LOCATION: 11..16
OTHER INFORMATION: /note= "spacer"
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= "pglu"
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoprolinone)"
US-08-968-466-2

Query Match      80.3%; Score 83.5; DB 3; Length 16;
Best Local Similarity 87.5%; Pred. No. 5.3e-05;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY      2 HWSYGLRPGSSPPPC 17
Db      2 HWSYGLRPG-RPPPC 16

RESULT      5
US-08-458-814-6
Sequence 6, Application US/08458814
Patent No. 6103243
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J
APPLICANT: DE AIZPURA, Henry J
APPLICANT: HOWE, Peter
APPLICANT: RAND, Keith N
TITLE OF INVENTION: ORAL VACCINES
```

```
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Larnet
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,814
FILING DATE: 02-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/327,822
FILING DATE: 18-OCT-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU86/00135
FILING DATE: 14-MAY-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH3104
FILING DATE: 25-OCT-1985
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH0566
FILING DATE: 15-MAY-1985
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/155/BIAU
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 904136
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 52 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-458-814-6

Query Match      62.5%; Score 65; DB 3; Length 52;
Best Local Similarity 83.3%; Pred. No. 0.037;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      2 HWSYGLRPGSSP 13
Db      37 HWSYGLRPGDP 48

RESULT      6
US-08-458-814-7
Sequence 7, Application US/08458814
Patent No. 6103243
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J
APPLICANT: DE AIZPURA, Henry J
APPLICANT: HOWE, Peter
APPLICANT: RAND, Keith N
TITLE OF INVENTION: ORAL VACCINES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Larnet
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,814
FILING DATE: 02-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/327,822
FILING DATE: 18-OCT-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU86/00135
FILING DATE: 14-MAY-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH3104
FILING DATE: 25-OCT-1985
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH0566
FILING DATE: 15-MAY-1985
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/155/BIAN
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 55 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-458-814-7

Query Match 62.5%; Score 65; DB 3; Length 55;
Best Local Similarity 83.3%; Pred. No. 0.039;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HMSGRLPGSSP 13
DB 39 HMSGRLPGSDP 50

RESULT 7
US-08-453-588-23
Sequence 23, Application US/08453588
Patent No. 5684145
GENERAL INFORMATION:
APPLICANT: Anna van der Zee, Irma Marianne van Die,
APPLICANT: Willem Pieter Martin Hoekstra,
APPLICANT: Josephus Theodorus Gieken.
TITLE OF INVENTION: Carrier system against GnRH
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Akzo No. 5684145el Patent Department
STREET: 1300 Piccard Drive, Suite 206
CITY: Rockville
STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,588
FILING DATE: 30-MAY-1995

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/078,661
FILING DATE: 16-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mary E. Gormley
REGISTRATION NUMBER: 34,409
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 258-5200
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-453-588-23

Query Match 59.6%; Score 62; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HMSGRLPGS 11
DB 3 HMSGRLPGS 12

RESULT 8
US-08-521-079-23
Sequence 23, Application US/08521079
Patent No. 6019983
GENERAL INFORMATION:
APPLICANT: Anna van der Zee, Irma Marianne van Die,
APPLICANT: Willem Pieter Martin Hoekstra,
APPLICANT: Josephus Theodorus Gieken.
TITLE OF INVENTION: Carrier system against GnRH
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Akzo Pharma
STREET: 1330 Piccard Drive
CITY: Rockville
STATE: Maryland
COUNTRY: U.S.A.
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/521,079
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/078,661
FILING DATE: 16-JUN-1993
APPLICATION NUMBER: EPA No. 6019983 92.201.775.1
FILING DATE: 18-JUN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: William M. Blackstone
REGISTRATION NUMBER: 29,772
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 258-5200
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-521-079-23

Query Match 59.6%; Score 62; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 11
Db 3 HWSYGLRPGS 12

RESULT 9
US-07-690-983D-18
Sequence 18, Application US/07690983D
Patent No. 5403586
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J.
APPLICANT: STEWART, Andrew G.
APPLICANT: TSONIS, Con G.
TITLE OF INVENTION: FUSION PROTEINS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/690,983D
FILING DATE: 25-JUN-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00373
FILING DATE: 24-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/148 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
INFORMATION FOR SRO ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-690-983D-18

Query Match 59.6%; Score 62; DB 1; Length 17;
Best Local Similarity 90.9%; Pred. No. 0.03;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 12
Db 5 HWSYGLRPGS 15

RESULT 10
US-07-690-983D-28
Sequence 28, Application US/07690983D
Patent No. 5403586
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J.
APPLICANT: STEWART, Andrew G.
APPLICANT: TSONIS, Con G.
TITLE OF INVENTION: FUSION PROTEINS
NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/690,983D
FILING DATE: 25-JUN-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00373
FILING DATE: 24-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/148 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
INFORMATION FOR SRO ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-690-983D-28

Query Match 59.6%; Score 62; DB 1; Length 18;
Best Local Similarity 90.9%; Pred. No. 0.032;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 12
Db 6 HWSYGLRPGS 16

RESULT 11
US-08-387-156-4
Sequence 4, Application US/08387156
Patent No. 5723129
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HUI P. A.
TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/387,156
FILING DATE: 10-FEB-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-387-156-4

Query Match 59.6%; Score 62; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 0.083;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGS 11
|||||

DB 2 HWSYGLRPGS 11

RESULT 12
US-08-694-865-4
Sequence 4, Application US/08694865
Patent No. 5837268
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANN, JOHN G.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,865
FILING DATE: 09-AUG-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-865-4

Query Match 59.6%; Score 62; DB 2; Length 49;
Best Local Similarity 100.0%; Pred. No. 0.083;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGS 11

DB 2 HWSYGLRPGS 11
|||||

RESULT 13
US-08-878-748-4
Sequence 4, Application US/08878748
Patent No. 5969126
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HOW P.A.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,748
FILING DATE: 19-JUN-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-748-4

Query Match 59.6%; Score 62; DB 2; Length 49;
Best Local Similarity 100.0%; Pred. No. 0.083;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGS 11
|||||

DB 2 HWSYGLRPGS 11

RESULT 14
US-09-124-491-4
Sequence 4, Application US/09124491
Patent No. 6022960
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANN, JOHN G.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:

ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,491
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491.4

Query Match 59.6%; Score 62; DB 3; Length 49;
Best Local Similarity 100.0%; Pred. No. 0.083;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGS 11
DB 2 HWSYGLRPGS 11

RESULT 15
US-08-387-156-10
Sequence 10, Application US/08387156
Patent No. 5723129
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/387,156
FILING DATE: 10-FEB-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-387-156-10

Query Match 59.6%; Score 62; DB 1; Length 544;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGS 11
DB 495 HWSYGLRPGS 504

Search completed: March 2, 2001, 10:54:36
Job time: 601 sec

THIS PAGE BLANK (USPTO)

<hr/>					
RESULT	1				
ID	O9PRHO				
AC	O9PRHO:	PRELIMINARY;	PRT:	91 AA.	
DT	01-MAY-2000 (TREMblrel. 13,	Created)			
DT	01-May-2000 (TREMblrel. 13,	Last sequence update)			
DT	01-JUN-2000 (TREMblrel. 14,	last annotation update)			
DE	PREPRO-MGNRH PRECURSOR.				
OC	Anquilla japonica (Japanese eel).				
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;				
OX	NCB1_TaxId=7937;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=BRAIN;				
RA	Okubo K., Suetake H., Aida K.;				
RT	"Expression of two gonadotropin-releasing hormone (GnRH) precursor				
RL	genes in various tissues of the Japanese eel and evolution of GnRH."				
LN	Zool. Sci. 16:471-478(1999).				
RP	[2]				
RA	SEQUENCE FROM N.A.				
RP	Okubo K., Suetake H., Aida K.;				
RT	A splicing variant for the prepro-mammalian gonadotropin-releasing				
RL	hormone (prepro-mGnRH) mRNA is present in the brain and various				
RT	peripheral tissues of the Japanese eel."				
RL	Zool. Sci. 16:645-651(1999).				
DR	EMBL; AB026989; BAA82608.1; -				
DR	EMBL; AB026991; BAA83597.1; -				
DR	INTERPRO: IPR002012; -				
DR	PFAM: PF00446; GNRH; 1.				
DR	PROSITE; PS00473; GNRH; 1.				
KW	Signal.				
FT	SIGNAL	1	22	POTENTIAL.	
FT	CHAIN	23	32	MGNRH.	
FT	CHAIN	33	91	GNRH ASSOCIATED PEPTIDE.	
SQ	SEQUENCE	91 AA; 9893 MW;		BALSC9DC08434A7B CRGC4;	
<hr/>					
Query Match	Similarity	55.8%;	Score 58;	DB 13;	Length 91;
Best Local	Similarity	100.0%;	Pred. NO. 0.22;		
Matches	9; Conservative	0;	Mismatches	0;	Indels
				0;	Gaps
					0;

OY 2 HWSYGLRPG 10
 |||||
 Db 24 HWSYGLRPG 32

RESULT 2
 ID 088020 PRELIMINARY; PRT; 112 AA.

AC 088020:
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, last annotation update)
 DE (MM316ZA) DNA, COMPLETE CDS.'S.
 GN VPX.
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
 NCBI_TaxID=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94016835; Pubmed=8411355;
 RA Kodama T., Mori K., Kawahara T., Ringler D.J., Destroiers R.C.;
 RT Analysis of simian immunodeficiency virus sequence variation in
 RL tissues of rhesus macaques with simian AIDS.*;
 J. Virol. 67:6522-6534(1993).
 DE EMBL: L22807; AAA47688.1; -;
 DR INTERPRO: IPRO00012; -;
 DR PRAM: PF000522; VPR: 1;
 DR PRINTS: PR00444; HIVVPRPX.
 SQ SEQUENCE 112 AA; 12945 MW; A933BF982AFD7BE8 CRC64;

Query Match 51.0%; Score 53; DB 12; Length 112;
 Best Local Similarity 60.0%; Pred. No. 1.4;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16
 | : | | | | | | | |
 Db 94 HGAEGWRRPGPPPP 108

RESULT 3
 ID 09Y126 PRELIMINARY; PRT; 87 AA.
 AC 09Y126:
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, last sequence update)
 DT 01-MAY-2000 (TREMblrel. 13, last annotation update)
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN)
 DE (FRAGMENT).
 OS Sparus aurata (Gilthead sea bream).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 OC Sparidae; Sparus.
 NCBI_TaxID=81175;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RA Nabissi M.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 FOLLICLE-STIMULATING HORMONES.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 DE EMBL: AF046801; AAD02427.1; -;
 DR INTERPRO: IPRO02012; -;
 DR PROSITE: PS00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KM Hormone; Amidation.
 FT NON_TER 1 87
 SQ SEQUENCE 87 AA; 9871 MW; 00246353D96782A CRC64;

Query Match 50.0%; Score 52; DB 13; Length 87;
 Best Local Similarity 88.9%; Pred. No. 1.5;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
 |||||
 Db 22 HWSYGLSPG 30

RESULT 4
 ID 073812 PRELIMINARY; PRT; 95 AA.

AC 073812:
 DT 01-AUG-1998 (TREMblrel. 07, Created)
 DT 01-AUG-1998 (TREMblrel. 07, last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, last annotation update)
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN).
 OS Morone saxatilis (Striped bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 OC Moronidae; Morone.
 NCBI_TaxID=34816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chow M.M., Kight K.E., Gotthilf Y., Alok D., Zohar Y.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 FOLLICLE-STIMULATING HORMONES.
 CC EMBL: AF056314; AAD03817.1; -;
 DR INTERPRO: IPRO02012; -;
 DR PRAM: PF00446; GNRH; 1.
 DR PROSITE: PS00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 DR PRODOM: PD00581; -; 1.
 KM Hormone; Amidation.
 SQ SEQUENCE 95 AA; 10411 MW; 980C6988FC279BFC CRC64;

Query Match 50.0%; Score 52; DB 13; Length 95;
 Best Local Similarity 88.9%; Pred. No. 1.7;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
 |||||
 Db 24 HWSYGLSPG 32

RESULT 5
 ID 09IA10 PRELIMINARY; PRT; 99 AA.
 AC 09IA10:
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, last annotation update)
 DE GONADOTROPIN RELEASING HORMONE SEABREAM ISOFORM.
 OS Dicentrarchus labrax (European sea bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 OC Moronidae; Dicentrarchus.
 NCBI_TaxID=13489;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RA Gonzalez-Martinez D., Madigou T., Zmora N., Anglade I., Zanny S.,
 RA Zohar Y., Elizur A., Munoz-Cueto J.A., Kah O.;
 RT Differential expression of three different prepro-GNRH
 RT (gonadotrophin-releasing hormone) messengers in the brain of the
 RT European sea bass (Dicentrarchus labrax).;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN:
RA Zmora N., Zohar Y., Eliazur A.:
RT "3 GnRH form in the seabass Dicentrarchus labrax."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF224279; AAF62898.1; -
SQ SEQUENCE 99 AA; 10758 MW; ECBAEEC3CC02904 CRC64;

Query Match 50.0%; Score 52; DB 13; Length 99;
Best Local Similarity 88.9%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
| | | | | | | |
DB 28 HWSYGLSPG 36

RESULT 6
ID 004273 PRELIMINARY; PRT: 112 AA.
AC 004273;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE VPX PROTEIN (X ORF PROTEIN).
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RA Benveniste R.E., Heidecker G., Greenwood J., Gonda M.A.;
RL Submitted (AUG-1990) to the EMBL/GenBank/DBJ databases.
RC -I- MISCELLANEOUS: THIS IS A CAPTIVE MACAQUE NEMESTRINA ISOLATE.
DR EMBL: M32741; AAA91933.1; -
DR INTERPRO: IPR000012; -
DR PFW: PF00522; VPR: 1.
DR PRINTS: PR00444; HIVPRVFX.
KW AIDS.
SQ SEQUENCE 112 AA; 12904 MW; 7E1564F45AFD7BEA CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;
Best Local Similarity 60.0%; Pred. No. 2;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16
| : | | | | | | | |
DB 94 HGAGGWRPGPPPPPP 108

RESULT 7
ID 088050 PRELIMINARY; PRT: 112 AA.
AC 088050;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE (MM316x) DNA, COMPLETE CDS.S.
GN VPX.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-94016835; PubMed-8411355;
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;
RT "Analysis of simian immunodeficiency virus sequence variation in
RT tissues of rhesus macaques with simian AIDS."
RL J. Virol. 67:6522-6534(1993).
DR EMBL: L22804; AAA47679.1; -
DR INTERPRO: IPR000012; -
DR PRINTS: PR00444; HIVPRVFX.
SQ SEQUENCE 112 AA; 12846 MW; 7FC007A3BAFD7BEA CRC64;

DR PFW: PF00522; VPR: 1.
DR PRINTS: PR00444; HIVPRVFX.
SQ SEQUENCE 112 AA; 12872 MW; 44BE094EF1955B60 CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;
Best Local Similarity 60.0%; Pred. No. 2;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16
| : | | | | | | | |
DB 94 HGAGGWRPGPPPPPP 108

RESULT 8
ID 088057 PRELIMINARY; PRT: 112 AA.
AC 088057;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE (MM316x) DNA, COMPLETE CDS.S.
GN VPX.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-94016835; PubMed-8411355;
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;
RT "Analysis of simian immunodeficiency virus sequence variation in
RT tissues of rhesus macaques with simian AIDS."
RL J. Virol. 67:6522-6534(1993).
DR EMBL: L22791; AAA72399.1; -
DR INTERPRO: IPR000012; -
DR PFW: PF00522; VPR: 1.
DR PRINTS: PR00444; HIVPRVFX.
SQ SEQUENCE 112 AA; 12842 MW; BDB78F982AFD7BF2 CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;
Best Local Similarity 60.0%; Pred. No. 2;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16
| : | | | | | | | |
DB 94 HGAGGWRPGPPPPPP 108

RESULT 9
ID 088058 PRELIMINARY; PRT: 112 AA.
AC 088058;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE (MM316L) DNA, COMPLETE CDS.S.
GN VPX.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-94016835; PubMed-8411355;
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;
RT "Analysis of simian immunodeficiency virus sequence variation in
RT tissues of rhesus macaques with simian AIDS."
RL J. Virol. 67:6522-6534(1993).
DR EMBL: L22792; AAA47643.1; -
DR INTERPRO: IPR000012; -
DR PFW: PF00522; VPR: 1.
DR PRINTS: PR00444; HIVPRVFX.
SQ SEQUENCE 112 AA; 12846 MW; 7FC007A3BAFD7BEA CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;
Best Local Similarity 60.0%; Pred. No. 2;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16
1 : 1 1 1 1 1 1 1
Db 94 HGAGGMRPGRP PPPPP 108

RESULT 10
ID 088064 PRELIMINARY; PRT; 112 AA.
AC 088064;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE (MM3160) DNA, COMPLETE CDS'S.
GN VPX.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94016835; PubMed=8411355;
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;
RT "Analysis of simian immunodeficiency virus sequence variation in
tissues of rhesus macaques with simian AIDS."
RL J. Virol. 67:6522-6534 (1993).
DR EMBL: L22795; AAA47652.1; -.
DR INTERPRO: IPR000012; -.
DR PFAM: PF00522; VPR; 1.
DR PRINTS: PR00444; HIVPRVPX.
SQ SEQUENCE 112 AA; 12935 MW; D32070A4DAC1CF5 CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;

Best Local Similarity 60.0%; Pred. No. 2;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16
1 : 1 1 1 1 1 1 1
Db 94 HGAGGMRPGRP PPPPP 108

RESULT 11
ID 088068 PRELIMINARY; PRT; 112 AA.
AC 088068;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE (MM3160) DNA, COMPLETE CDS'S.
GN VPX.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94016835; PubMed=8411355;
RA Kodama T., Mori K., Kawahara T., Ringler D.J., Desrosiers R.C.;
RT "Analysis of simian immunodeficiency virus sequence variation in
tissues of rhesus macaques with simian AIDS."
RL J. Virol. 67:6522-6534 (1993).
DR EMBL: L22797; AAA47658.1; -.
DR INTERPRO: IPR000012; -.
DR PFAM: PF00522; VPR; 1.
DR PRINTS: PR00444; HIVPRVPX.
SQ SEQUENCE 112 AA; 12841 MW; 8B82EF96CA137BF0 CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;

Best Local Similarity 60.0%; Pred. No. 2;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16
1 : 1 1 1 1 1 1 1
Db 94 HGAGGMRPGRP PPPPP 108

RESULT 12
ID 088125 PRELIMINARY; PRT; 112 AA.
AC 088125;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE VPX PROTEIN.
GN VPX.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90272009; PubMed=1971917;
RA Dewhurst S., Embretson J.E., Anderson D.C., Mullins J.I., Fultz P.N.;
RT "Sequence analysis and acute pathogenicity of molecularly cloned
SIVSM-PB14."
RL Nature 345:636-640 (1990).
DR EMBL: L03296; AAA47763.1; -.
DR INTERPRO: IPR000012; -.
DR PFAM: PF00522; VPR; 1.
DR PRINTS: PR00444; HIVPRVPX.
SQ SEQUENCE 112 AA; 12877 MW; BEADC7FE6CFD2EAC CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;

Best Local Similarity 60.0%; Pred. No. 2;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPP 16
1 : 1 1 1 1 1 1 1
Db 94 HGAGGMRPGRP PPPPP 108

RESULT 13
ID 088131 PRELIMINARY; PRT; 112 AA.
AC 088131;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE VPX PROTEIN.
GN VPX.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90272009; PubMed=1971917;
RA Dewhurst S., Embretson J.E., Anderson D.C., Mullins J.I., Fultz P.N.;
RT "Sequence analysis and acute pathogenicity of molecularly cloned
SIVSM-PB14."
RL Nature 345:636-640 (1990).
DR EMBL: L03297; AAA47771.1; -.
DR INTERPRO: IPR000012; -.
DR PFAM: PF00522; VPR; 1.
DR PRINTS: PR00444; HIVPRVPX.
SQ SEQUENCE 112 AA; 12851 MW; 9C3EC7FE6CFD3D14 CRC64;

Query Match 50.0%; Score 52; DB 12; Length 112;
Best Local Similarity 60.0%; Pred. No. 2;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPP 16
 I : I I I I I I I
 DB 94 HGAGGWRPGPPPPP 108

RESULT 14

O88137 PRELIMINARY; PRT: 112 AA.
 AC O88137:
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE VPX PROTEIN.
 GN VPX.
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90272009; PubMed=1971917;
 RA Dewhurst S., Embretson J.E., Anderson D.C., Mullins J.I., Fultz P.N.;
 RT "Sequence analysis and acute pathogenicity of molecularly cloned
 RT SIVSM-PBJ14."
 RL Nature 345:636-640(1990).
 DR EMBL: L03295; AAB59772.1; -
 DR INTERPRO: IPR000012; -
 DR PRAM: PF005522; VPR: 1.
 DR PRINTS: PR00444; HIVPRVPX.
 SQ SEQUENCE 112 AA; 12881 MW; C36D1AE2A0FD270E CRC64;

Query Match

Best Local Similarity 50.0%; Score 52; DB 12; Length 112;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPP 16
 I : I I I I I I I
 DB 94 HGAGGWRPGPPPPP 108

RESULT 15

O88145 PRELIMINARY; PRT: 112 AA.
 AC O88145:
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE VPX PROTEIN.
 GN VPX.
 OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92124755; PubMed=1733112;
 RA Novembre F.J., Hirsch V.M., McClure H.M., Fultz P.N., Johnson P.R.;
 RT "SIV from stump-tailed macaques: molecular characterization of a
 RT highly transmissible primate lentivirus."
 RL Virology 186:783-787(1992).
 DR EMBL: M83293; AAA91942.1; -
 DR INTERPRO: IPR000012; -
 DR PRAM: PF005522; VPR: 1.
 DR PRINTS: PR00444; HIVPRVPX.
 SQ SEQUENCE 112 AA; 12934 MW; ADD7F9DC45868440 CRC64;

Query Match

Best Local Similarity 50.0%; Score 52; DB 12; Length 112;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPP 16
 I : I I I I I I I
 DB 94 HGAGGWRPGPPPPP 108

Search completed: March 2, 2001, 11:00:48
 Job time: 458 sec

THIS PAGE BLANK (USPTO)

FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).
 FT NON_TER 61 61
 SQ SEQUENCE 61 AA: 6828 MW: 63962A1AE319B8F0 CRC64:

Query Match 55.8%; Score 58; DB 1; Length 61;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
 Db 2 HWSYGLRPG 10

RESULT 2

CON1_MESAU STANDARD; PRT; 63 AA.
 ID CON1_MESAU
 AC 009163;
 DT 15-DEC-1998 (rel. 37, Created)
 DT 15-DEC-1998 (rel. 37, Last sequence update)
 DT 30-MAY-2000 (rel. 39, Last annotation update)
 DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
 (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
 HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]
 DE (FRAGMENT).
 GN GNRH1 OR GNRH OR LHRH.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jansen H.T., Stevens P.J., Zeidler P., Lehman M.N.;
 Submitted (MAR-1987) to the EMBL/GenBank/DDI databases.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 HORMONES.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----

CC EMBL: U91938; AAB51302.1; -
 DR INTERPRO: IPR002012; -
 DR PFAM: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta.
 FT NON_TER 1 1
 FT CHAIN 1 >63
 FT PEPTIDE 1 10
 FT PEPTIDE 14 >63
 FT ACT_SITE 3 3
 FT MOD_RES 1 1
 FT MOD_RES 10 10
 FT MOD_RES 10 10
 FT NON_TER 63 63
 FT SEQUENCE 63 AA: 7370 MW: FC94995676F77180 CRC64;

Query Match 55.8%; Score 58; DB 1; Length 63;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
 Db 2 HWSYGLRPG 10

RESULT 3

CON1_MACMU STANDARD; PRT; 67 AA.
 ID CON1_MACMU
 AC P55247;
 DT 01-OCT-1996 (rel. 34, Created)
 DT 01-OCT-1996 (rel. 34, Last sequence update)
 DT 30-MAY-2000 (rel. 39, Last annotation update)
 DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
 (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
 HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]
 DE (FRAGMENT).
 GN GNRH1 OR GNRH OR LHRH.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-HYPOTHALAMUS;
 RC MEDLINE-95124501; PubMed-7545971;
 RA Ma Y.J., Costa M.E., Ojeda S.R.;
 RT "Developmental expression of the genes encoding transforming growth
 factor alpha and its receptor in the hypothalamus of female rhesus
 macaques."
 RT Neuroendocrinology 60:346-359(1994).
 RL
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 HORMONES.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----

CC EMBL: S75918; AAB33096.1; -
 DR INTERPRO: IPR002012; -
 DR PFAM: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Signal.
 FT NON_TER 1 1
 FT SIGNAL 1 5
 FT CHAIN 6 >67
 FT PEPTIDE 6 15
 FT PEPTIDE 19 >67
 FT ACT_SITE 8 8
 FT MOD_RES 6 6
 FT MOD_RES 15 15
 FT MOD_RES 15 15
 FT NON_TER 67 67
 FT SEQUENCE 67 AA: 7573 MW: 505394DAA261A3F2 CRC64;

Query Match 55.8%; Score 58; DB 1; Length 67;
 Best Local Similarity 100.0%; Pred. No. 0.17;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
 Db 7 HWSYGLRPG 15

```

RESULT 4
GON1_XENLA STANDARD: PRT: 89 AA.
AC P45656.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE (LH-RH) (LULIBERIN I).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-FOREBRAIN;
RA MEDLINE=9418563; PubMed=8137750;
RA Hayes W.P., Wray S., Battey J.F.;
RT "The frog gonadotropin-releasing hormone-I (GNRH-I) gene has a
RT mammalian-like expression pattern and conserved domains in
RT GNRH-associated peptide, but brain onset is delayed until
RT metamorphosis."
RL Endocrinology 134:1835-1844(1994).
CC -I- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -I- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L28040; AAA9728.1; .
DR INTERPRO; IPR002012; .
DR PFAM; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 23
FT CHAIN 24 89 PROGONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT CHAIN 37 89 GONADOTROPIN-RELEASING HORMONE ASSOCIATED
PEPTIDE.
FT PEPTIDE 37 85 GNRH-ASSOCIATED PEPTIDE I (GAP).
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SO SEQUENCE 89 AA; 10246 MW; 6FAF36FBAE0D4284 CRC64;

Query Match 55.8%; Score 58; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33

RESULT 5
GON1_MOUSE STANDARD: PRT: 90 AA.
AC P13562.
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR
DE I].
OS GNRH1 OR GNRH.
GN Mus musculus (Mouse).

```

```

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87069928; PubMed=3024317;
RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
RA Phillips H.S., Nikolic K., Seeburg P.H.;
RT "A deletion truncating the gonadotropin-releasing hormone gene is
RT responsible for hypogonadism in the hpg mouse."
RL Science 234:1366-1371(1986).
CC -I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -----
CC -I- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collabora
CC between the Swiss Institute of Bioinformatics and the EMBL outstation i
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14872; AAA37717.1; .
DR MED: MG1:95789; GNRH.
DR INTERPRO; IPR002012; .
DR PFAM; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 21
FT CHAIN 22 90 PROGONADOLIBERIN I.
FT PEPTIDE 22 31 GONADOLIBERIN I.
FT PEPTIDE 35 90 PROLACTIN RELEASE-INHIBITING FACTOR I.
FT ACT_SITE 24 24 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
ACTIVITY.
FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).
SO SEQUENCE 90 AA; 10337 MW; 1C0766FA4826E4D9 CRC64;

Query Match 55.8%; Score 58; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
DB 23 HWSYGLRPG 31

RESULT 6
GON1_PIG STANDARD: PRT: 91 AA.
AC P49921.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-HYPOTHALAMUS;
RA Weesner G.D., Mattern R.L., Becker B.A.;
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 24-33.
RA MEDLINE=72114303; PubMed=4946067;
RA Baba Y., Matsuo H., Schally A.V.;

```

```

RT "Structure of the porcine LH- and FSH-releasing hormone. II.
RT Confirmation of the proposed structure by conventional sequential
RT analyses."
RL Biochem. Biophys. Res. Commun. 44:459-463(1971).
RN [3]
RP SYNTHESIS OF GONADOLIBERIN.
RX MEDLINE-72065376; PubMed-4942726;
RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
RT phase method."
RL Biochem. Biophys. Res. Commun. 45:822-827(1971).
RN [4]
RP SYNTHESIS OF GONADOLIBERIN.
RX MEDLINE-72117544; PubMed-4946275;
RA Baba Y., Arimura A., Schally A.V.;
RT "On the tryptophan residue in porcine LH and FSH-releasing hormone."
RL Biochem. Biophys. Res. Commun. 45:483-487(1971).
RN [1]
RP FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
RP THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
RP HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L32864; AAA31066.1; -.
CC DR PIR: A01411; RHFG.
CC DR INTERPRO: IPR002012; -.
CC PFAM: PF00446; GNRH; 1.
CC PROSITE: PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Placenta; Signal.
CC FT SIGNAL 1 23
CC FT CHAIN 24 91
CC FT PEPTIDE 24 33
CC FT PEPTIDE 34 91
CC FT ACT_SITE 26 26
CC FT MOD_RES 24 24
CC FT MOD_RES 33 33
CC FT SEQUENCE 91 AA; 10090 MW; 8340474P32DDAA9 CRC64;
SQ
Query Match 55.8%; Score 58; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33
RESULT 7
GON1_HUMAN STANDARD: PRT; 92 AA.
AC P01148;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I) (GONADORELIN); GNRH-ASSOCIATED
DE PEPTIDE I].
DE GNRH1 OR GNRH OR LHRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE-8936682; PubMed-2671939;
RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT "The complete nucleotide sequence of the human gonadotropin-releasing
RT hormone gene."
RL Nucleic Acids Res. 17:6403-6403(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-86094338; PubMed-2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat."
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-85012739; PubMed-6090951;
RA Seeburg P.H., Adelman J.P.;
RT "Characterization of cDNA for precursor of human luteinizing hormone
RT releasing hormone."
RL Nature 311:666-668(1984).
RN [4]
RP SEQUENCE OF 24-33.
RX MEDLINE-83126573; PubMed-6760865;
RA Tan L., Rousseau P.;
RT "The chemical identity of the immunoreactive LHRH-like peptide
RT biosynthesized in the human placenta."
RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
RN [1]
RP FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
RP THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
RP HORMONES.
CC -1- PHARMACEUTICAL: AVAILABLE UNDER THE NAMES FACTREL (AYERST LABS),
CC LUTEREPUL OR LUTERELF (FERRING PHARMACEUTICALS) AND RELISORN
CC (SERONO).
CC -----
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X01059; CAA25526.1; -.
CC DR EMBL: M12578; AAA35916.1; -.
CC DR PIR: X15215; CAA33285.1; -.
CC DR PIR: A01410; RHFG.
CC DR PIR: A26173; A26173.
CC DR PIR: S05308; S05308.
CC MIM: 152760; -.
CC DR INTERPRO: IPR002012; -.
CC PFAM: PF00446; GNRH; 1.
CC PROSITE: PS00473; GNRH; 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Placenta; Pharmaceutical; Signal.
CC FT SIGNAL 1 23
CC FT CHAIN 24 92
CC FT PEPTIDE 24 33
CC FT PEPTIDE 37 92
CC FT ACT_SITE 26 26
CC FT MOD_RES 24 24
CC FT MOD_RES 33 33
CC FT CONFLICT 16 16
CC FT SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;
SQ
Query Match 55.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG 10

```

Db 25 HWSYGLRPG 33

```

RESULT 8
GONL_RAT STANDARD: PRT: 92 AA.
ID GONL_RAT
AC P07490:
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR (CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR
DE I).
GN GNRH1 OR GNRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat."
RT Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RL [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=89384661; PubMed=2476669;
RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT "The rat gonadotropin-releasing hormone: SH locus: structure and
RT hypothalamic expression."
RT Mol. Endocrinol. 3:1257-1262(1989).
RL [3]
RN SEQUENCE FROM N.A.
RP TISSUE-THYMUS.
RC MEDLINE=93105480; PubMed=1468115;
RA Maier C.C., Marchetti B., Lebeuf R.D., Blalock J.E.;
RT "Thymocytes express a mRNA that is identical to hypothalamic
RT luteinizing hormone-releasing hormone mRNA."
RL Cell. Mol. Neurobiol. 12:447-454(1992).
RN [4]
RN SEQUENCE OF 1-47 FROM N.A.
RC TISSUE-HEART;
RA MEDLINE=87149087; PubMed=3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
RT DNA locus."
RL Science 235:1514-1517(1987).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S50870; AAB24572.1; -
CC EMBL; M12579; AAA41263.1; -
CC EMBL; M31670; AAA41264.1; -
CC EMBL; M15527; AAA42141.1; ALT_SEQ.
CC EMBL; M15529; AAA42139.1; -
CC EMBL; M15528; -; NOT_ANNOTATED_CDS.
CC PIR; B26173; RHRTG.
CC PIR; A48410; A48410.
CC INTERPRO; IPR002012; -

```

```

DR PFAM; PF00446; GNRH. 1.
DR PROSITE; PS00473; GNRH. 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 32
FT ACT_SITE 26 26
FT MOD_RES 24 24
FT MOD_RES 33 33
SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;

Query Match 55.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db 25 HWSYGLRPG 33

RESULT 9
GONL_TUPGB STANDARD: PRT: 92 AA.
ID GONL_TUPGB
AC Q95335;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR (CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
GN GNRH1 OR GNRH.
OS Tupia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupalidae; Tupala.
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE-HYPOTHALAMUS;
RC MEDLINE=97079639; PubMed=8921350;
RA Kaestlen T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGNRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GNRH gene expression in a
RT placental mammal."
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U63326; AAB16837.1; -
CC INTERPRO; IPR002012; -
CC PFAM; PF00446; GNRH. 1.
DR PROSITE; PS00473; GNRH. 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92
FT PEPTIDE 24 33
FT PEPTIDE 37 92
FT ACT_SITE 26 26
BY SIMILARITY.
PROGONADOLIBERIN I.
GONADOLIBERIN I.
GNRH-ASSOCIATED PEPTIDE I.
APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
ACTIVITY.

```

FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
 FT MOD_RES 33 33 SIMILARITY).
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
 FT SEQUENCE 92 AA; 10197 MW; 4FDBE2C58CF5F63B CRC64;
 SIMILARITY).

Query Match 55.8%; Score 58; DB 1; Length 92;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 Db 25 HWSYGLRPG 33

RESULT 10
 GONI_ALAMI STANDARD; PRT; 10 AA.
 AC P37041; P20407;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GONADOLIBERIN I (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I) (LH-RH I)
 DE (LULIBERIN I).
 OS Alligator mississippiensis (American alligator).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
 RN [1]
 RP SEQUENCE.
 RC TISSUE-BRAIN;
 RX MEDLINE=91352338; PubMed=1882082;
 RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
 Lance V., Stanson P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure of two forms of gonadotropin-releasing hormone
 from brains of the American alligator (Alligator mississippiensis).";
 RL Regul. Pept. 33:105-116(1991).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 DR INTERPRO: IPR002012;
 DR PIR: A60066; RHA01.
 DR PRAM: PF00446; GNRH: 1.
 DR PROSITE: PS00473; GNRH: 1.
 KW Hormone; Amidation; Hypothalamus.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1172 MW; 284B23D7286B45A3 CRC64;

Query Match 51.9%; Score 54; DB 1; Length 10;
 Best Local Similarity 88.9%; Pred. No. 0.085;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 Db 2 HWSYGLRPG 10

RESULT 11
 GONI_CHICK STANDARD; PRT; 92 AA.
 AC P37042; P20407;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
 DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
 DE (HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-WHITE LEHORN;
 RX MEDLINE=94059355; PubMed=7902095;
 RA Dunn I.C., Chen Y., Hook C., Sharp P.J., Sang H.M.;
 RT "Characterization of the chicken preprogonadotropin-releasing
 RT hormone-I gene.";
 RL J. Mol. Endocrinol. 11:19-29(1993).
 RN [2]
 RP SEQUENCE OF 24-33.
 RC TISSUE-HYPOTHALAMUS;
 RX MEDLINE=82265778; PubMed=7050119;
 RA King J.A., Millar R.P.;
 RT "Structure of chicken hypothalamic luteinizing hormone-releasing
 RT hormone. II. Isolation and characterization.";
 RL J. Biol. Chem. 257:10729-10732(1982).
 RN [3]
 RP SEQUENCE OF 24-33.
 RC TISSUE-HYPOTHALAMUS;
 RA King J.A., Millar R.P.;
 RT "Structure of avian hypothalamic gonadotropin-releasing hormone.";
 RL S. Afr. J. Sci. 78:124-125(1982).
 RN [4]
 RP SYNTHESIS OF 24-33.
 RX MEDLINE=82265777; PubMed=7050118;
 RA King J.A., Millar R.P.;
 RT "Structure of chicken hypothalamic luteinizing hormone-releasing
 RT hormone. I. Structural determination on partially purified
 RT material.";
 RL J. Biol. Chem. 257:10722-10728(1982).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X69491; CAA49246.1;
 DR PTR: S33507; S33507.
 DR INTERPRO: IPR002012;
 DR PRAM: PF00446; GNRH: 1.
 DR PROSITE: PS00473; GNRH: 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 FT SIGNAL. 1 23
 FT CHAIN 24 92
 FT PEPTIDE 24 33
 FT PEPTIDE 37 92
 FT MOD_RES 24 24 GNRH-ASSOCIATED PEPTIDE I.
 FT MOD_RES 33 33 PYRROLIDONE CARBOXYLIC ACID.
 SQ SEQUENCE 92 AA; 10206 MW; 61AEB7EBAF508B6A CRC64;

Query Match 51.9%; Score 54; DB 1; Length 92;
 Best Local Similarity 88.9%; Pred. No. 0.75;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
 Db 25 HWSYGLRPG 33

RESULT 12
 GONI_HAPBU STANDARD; PRT; 94 AA.
 AC P51918; 093387;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)

```

DE (LH-RH I) (LULIBERIN I).
GN *GNRH1.
OS Haplochromis butoni.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidel;
OC Clichidae; Astatotilapia.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95396797; PubMed-7667296;
RA White S.A., Kaslen T.L., Bond C.T., Adelman J.P., Fernald R.D.;
RT "Three gonadotropin-releasing hormone genes in one organism suggest
RT novel roles for an ancient peptide."
RL Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99061842; PubMed-9843638;
RA White R.B., Fernald R.D.;
RT "Ontogeny of gonadotropin-releasing hormone (GNRH) gene expression
RT reveals a distinct origin for GNRH-containing neurons in the
RT midbrain."
RL Gen. Comp. Endocrinol. 112:322-329(1998).
RN [3]
RP SEQUENCE OF 23-32.
RC TISSUE-PITUITARY;
RX MEDLINE-95372591; PubMed-7644702;
RA Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,
RA White S.A., Francis R.C., Fernald R.D., Licht P., Wardy C.,
RA Sherwood N.M.;
RT "Primary structure of solitary form of gonadotropin-releasing hormone
RT (GNRH) in cichlid pituitary; three forms of GNRH in brain of cichlid
RT and pumpkinseed fish."
RL Regul. Pept. 57:43-53(1995).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE
CC RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALAMIC-PITUITARY-
CC GONADAL AXIS.
CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS
CC TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPOPHYSAL AXONS.
CC -1- MASS SPECTROMETRY: MW-1113.9; METHOD-MALDI; RANGE-23-32.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U31865; AAC59691.1; -
DR EMBL: AF076961; AAC27716.1; -
DR INTERPRO: IPR002012; -
DR PFMW: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KM Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Multigene family.
FT SIGNAL 1 22
FT CHAIN 23 94
FT PEPTIDE 23 32
FT MOD_RES 36 94
FT MOD_RES 23 23
FT MOD_RES 32 32
FT CONFLICT 86 94
SO SEQUENCE 94 AA; 10362 MW; E57DBA8333278D7 CRC64;

Query Match 50.0%; Score 52; DB 1; Length 94;
Best Local Similarity 88.9%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 HWSYGLRPG 10
DB 24 HWSYGLSPG 32

```

```

RESULT 13
GNL_PAGMA STANDARD; PRT; 95 AA.
ID GNL_PAGMA
AC P70074;
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE (LH-RH I) (LULIBERIN I).
OS Pagrus major (Red sea bream) (Chrysophrys major).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Chrysophrys.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Okuzawa K., Graneman J., Bogerd J., Goos H., Zohar Y., Kagawa H.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D86582; BAA13129.1; -
DR INTERPRO: IPR002012; -
DR PFMW: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KM Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Multigene family.
FT SIGNAL 1 23
FT CHAIN 24 95
FT PEPTIDE 24 33
FT MOD_RES 37 95
FT MOD_RES 24 24
FT MOD_RES 33 33
SO SEQUENCE 95 AA; 10566 MW; 61E79C990328D73E CRC64;

Query Match 50.0%; Score 52; DB 1; Length 95;
Best Local Similarity 88.9%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 HWSYGLRPG 10
DB 25 HWSYGLSPG 33

RESULT 14
GNL_PAGMA STANDARD; PRT; 95 AA.
ID GNL_PAGMA
AC P51919;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE (LH-RH I) (LULIBERIN I) (SBGNRH).
OS Sparus aurata (Gilthead sea bream).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Sparus.

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=95268499; PubMed-7749463;
RA Gotlib Y., Elizur A., Chow M., Chen T.T., Zohar Y.;
RT "Molecular cloning and characterization of a novel gonadotropin-
RL releasing hormone from the gilthead seabream (Sparus aurata).";
RM Mol. Mar. Biol. Biotechnol. 4:27-35(1995).
RN [2]
RP SEQUENCE OF 26-35.
RC TISSUE=BRAIN;
RX MEDLINE=95083645; PubMed-7991588;
RA Powell J.F.F., Zohar Y., Elizur A., Park M., Fischer W.H.,
RA Craig A.G., Rivier J.E., Lovejoy D.A., Sherwood N.M.;
RT "Three forms of gonadotropin-releasing hormone characterized from
RL brains of one species.";
RM Proc. Natl. Acad. Sci. U.S.A. 91:12081-12085(1994).
RN [3]
RP FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
RC -1- MASS SPECTROMETRY: MW=1113.6; METHOD=MALDI; RANGE=26-35.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U30320; AAA75469.1; -
DR INTERPRO: IPR002012; -
DR PFAM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KV Cleavage on pair of basic residues: Hormone; Amidation; Hypothalamus;
KW Signal, Multigene family.
FT SIGNAL 1 25
FT CHAIN 1 25
FT PEPTIDE 26 35
FT PEPTIDE 26 35
FT MOD_RES 26 26
FT MOD_RES 35 35
SQ SEQUENCE 95 AA; 10753 MW; 49313FDE6B87DA CRC64;

Query Match
Best Local Similarity 50.0%; Score 52; DB 1; Length 95;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
Db 27 HWSYGLSPG 35

RESULT 15
VPX_STVM1 STANDARD; PRT; 112 AA.
AC P05917;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-OCT-1989 (Rel. 12, Last annotation update)
DE VPX PROTEIN (X ORF PROTEIN).
GN VPX.
OS Simian immunodeficiency virus (Mm142-83 isolate) (STV-MAC).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87287230; PubMed=3649576;
RA Chakrabarti L., Guyader M., Alizon M., Daniel M.D., Desrosiers R.C.,
RA Hjalals P., Sonigo P.;
RT "Sequence of simian immunodeficiency virus from macaque and its
RT relationship to other human and simian retroviruses.";
RL Nature 328:543-547(1987).
CC -1- MISCELLANEOUS: THIS IS A MACAQUE ISOLATE.

```

```

CC [1]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: Y00277; CAA68382.1; -
DR PIR: D28887; ASLTX3.
DR HIV: M16403; VPXMM142.
DR INTERPRO: IPR000012; -
DR PFAM: PF00522; VPR; 1.
DR PRINTS: PR00444; HIVVPRVPX.
KW AIDS.
SQ SEQUENCE 112 AA; 12906 MW; 46F564F45AFPD960 CRC64;

Query Match
Best Local Similarity 60.0%; Score 52; DB 1; Length 112;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSPPP 16
Db 94 HGAGGWRPGPPPPP 108

```

Search completed: March 2, 2001, 11:01:20
Job time: 435 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:55:39 ; Search time 56.87 Seconds
(without alignments)
20.297 Million cell updates/sec

Title: US-09-306-689-4

Perfect score: 104

Sequence: 1 XHMSYGLRPGSSPPPC 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	55.8	10	1	RHPCG gonadoliberin - pi
2	58	55.8	10	1	RHSHG gonadoliberin - sh
3	58	55.8	67	2	I78541 gonadoliberin prec
4	58	55.8	89	2	I51423 gonadoliberin prec
5	58	55.8	90	1	RHMSG gonadoliberin prec
6	58	55.8	92	1	RHHUG gonadoliberin prec
7	58	55.8	92	1	RHRFG gonadoliberin prec
8	54	51.9	10	1	RHAQ1 gonadoliberin I -
9	54	51.9	92	2	I50644 gonadoliberin I pr
10	52	50.0	98	2	I50739 gonadotropin-relea
11	52	50.0	106	2	A30996 orf-X protein - si
12	52	50.0	112	1	ASLJX3 vpu protein - simi
13	52	50.0	112	2	S03066 gene X protein - h
14	51	49.0	112	1	ASLJST vpu protein - simi
15	51	49.0	862	2	T46289 hypothetical prote
16	50	48.1	112	1	ASLJX2 vpu protein - huma
17	50	48.1	112	1	ASLJX2 vpu protein - huma
18	50	48.1	112	2	SI2155 vpx protein - huma
19	50	48.1	112	2	SI3094 vpx protein - huma
20	49	47.1	775	2	E83400 pyroglutaminolase q
21	49	47.1	2129	1	RHJDS hypothetical prote
22	48	46.2	80	1	RHJDS gonadoliberin I pr
23	48	46.2	300	2	S41171 transrepressor pro
24	48	46.2	432	2	A25483 env polypeptide, r
25	48	46.2	676	1	VCMPV env polypeptide, r
26	48	46.2	833	2	G75621 arylesterase/monox
27	48	46.2	969	2	T15446 hypothetical prote
28	48	46.2	1091	2	T13170 diaphanous protein
29	47	45.2	199	2	SI4981 extensin class I (

30	47	45.2	316	2	A59021 aldehyde reductase
31	47	45.2	320	2	T48188 aldose reductase-1
32	47	45.2	367	2	T01751 gibberellin 20-oxi
33	47	45.2	1438	2	B71610 WD40 WEB-1 homolog
34	46	44.2	30	2	S59482 hydroxyproline-ric
35	46	44.2	157	2	T02034 early light-induce
36	46	44.2	181	2	T35751 hypothetical prote
37	46	44.2	256	2	S74928 hypothetical prote
38	46	44.2	309	1	S09257 homeotic protein H
39	46	44.2	422	2	T09742 drought-induced pr
40	46	44.2	546	2	T02029 DNA-binding protei
41	46	44.2	1373	2	JE0095 gastric mucin MUC5
42	45	43.3	10	2	A21114 gonadoliberin - ch
43	45	43.3	74	2	T51092 gonadotropin relea
44	45	43.3	82	2	I51180 gonadotropin-relea
45	45	43.3	82	2	I51365 gonadotropin-relea

ALIGNMENTS

```

RESULT 1
RHPGC
gonadoliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411
R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of
A:Reference number: A90172; MUID:72114303
A:Accession: A01411
A:Molecule type: protein
A:Residues: 1-10 <BAB>
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me
A:Reference number: A90176; MUID:72065376
A:Contents: annotation; synthesis
A>Note: the synthetic and natural hormones have the same physicochemical and biologic
R:Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A:Reference number: A90175; MUID:72117544
A:Contents: annotation
A>Note: Trp-3 appears to be essential for biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 55.8%; Score 58; DB 1; Length 10;
Best local similarity 100.0%; Pred. No. 0.054;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HMSYGLRPG 10
DB 2 HMSYGLRPG 10

RESULT 2
RHSHG
gonadoliberin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C:Accession: A93780; A01411
R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa
A:Reference number: A93780; MUID:72094314
A:Accession: A93780

```

A:Molecule type: protein
A:Residues: 1-10 <BUR>
A>Note: the natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormone
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 55.8%; Score 58; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
DB 2 HWSYGLRPG 10

RESULT 3
178541

gonadoliberin precursor - rhesus macaque (fragment)
N:Alternate names: luteinizing hormone releasing hormone
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C:Accession: I78541
R:Ma, Y.J.; Costa, M.E.; Ojeda, S.R.
Neuroendocrinology 60, 346-359, 1994
A>Title: Developmental expression of the genes encoding transforming growth factor alpha
A:Reference number: I58134; M0ID:95124501
A:Accession: I78541
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-67 <RES>
A:Cross-references: GB:S75918; NID:g912831; PIDN:AAB3096.1; PID:g912832
C:Superfamily: gonadoliberin

Query Match 55.8%; Score 58; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
DB 7 HWSYGLRPG 15

RESULT 4

gonadoliberin precursor - African clawed frog
N:Alternate names: luteinizing hormone releasing hormone
C:Species: Xenopus laevis (African clawed frog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: I51423
R:Hayes, W.P.; Wray, S.; Batley, J.F.
Endocrinology 134, 1835-1845, 1994
A>Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved domain
A:Reference number: I51423; M0ID:94185563
A:Accession: I51423
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-89 <HAY>
A:Cross-references: GB:I28040; NID:g496291; PIDN:AAA9728.1; PID:g496292
C:Genetics:
A:Gene: GnRH-I
C:Superfamily: gonadoliberin

Query Match 55.8%; Score 58; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10

DB 25 HWSYGLRPG 33

RESULT 5

RHMSG
gonadoliberin precursor - mouse
N:Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing hormone (LHRH)
C:Species: Mus musculus (house mouse)
C>Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
C:Accession: A47578
R:Masson, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikoli
Science 234, 1366-1371, 1986
A>Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible
A:Reference number: A47578; M0ID:87069928
A:Accession: A47578
A:Molecule type: DNA
A:Residues: 1-90 <MAS>
A:Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387475

C:Genetics:
A:Insertions: 45/3; 77/3

C:Function:
A:Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitrop
A>Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:22-31/Product: gonadoliberin #status predicted <GLS>
F:25-90/Product: gonadoliberin-associated protein #status predicted <GAP>
F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 55.8%; Score 58; DB 1; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
DB 23 HWSYGLRPG 31

RESULT 6
RHMSG

gonadoliberin precursor - human
N:Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasing hormone (LHRH)
C:Species: Homo sapiens (man)
C>Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 18-Jun-1999
C:Accession: S05308; A26173; A93342; A90108; A01410; S45718
R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
Nucleic Acids Res. 17, 6403-6404, 1989
A>Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone
A:Reference number: S05308; M0ID:8936682
A:Accession: S05308
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-92 <HAY>
A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A>Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone
A:Reference number: A94090; M0ID:86094338
A:Accession: A26173
A:Molecule type: mRNA
A:Residues: 1-92 <ADE>
A:Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749
R:Seeburg, P.H.; Adelman, J.P.
Nature 311, 666-668, 1984
A>Title: Characterization of cDNA for precursor of human luteinizing hormone releasing hormone
A:Reference number: A93342; M0ID:85012739

A:Accession: A93342
A:Molecule type: mRNA
A:Residues: 1-15, 'S', 17-92 <SEE>
A:Cross-references: GB:X01059; NID:934356; PIDN:CAA25526.1; PID:934357
A:Experimental source: placenta
R:Tan, L.; Rousseau, P.
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982
A>Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in
A:Reference number: A90108; MUID:83126573
A:Accession: A90108
A:Molecule type: protein
A:Residues: 24-33 <TAN>
A:Experimental source: placental trophoblasts
R:Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baummeister, W.; Amsterda
FBS Lett. 346, 203-206, 1994
A>Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by th
A:Reference number: S45718; MUID:94283597
A:Contents: annotation: degradation pathway of synthetic hormone
C:Genetics:
A:Gene: GDB:GNRH: LHRH; GRH
A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760
A:Map position: 8p21-8p11.2
A:Introns: 47/3; 79/3
C:Function:
A:Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitropin
A>Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-92/Product: progonaoliberin #status predicted <PGN>
F:24-33/Product: gonadoliberin #status experimental <MAT>
F:37-92/Product: gonadoliberin-associated protein #status predicted <GAP>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 55.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33

RESULT 7
RHRTG
gonadoliberin precursor - rat
N:Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hormo
N:Contents: gonadoliberin; prolactin release-inhibiting factor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
C:Accession: A40147; B26173; A48410
R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
Mol. Endocrinol. 3, 1257-1262, 1989
A>Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex
A:Reference number: A40147; MUID:89384661
A:Accession: A40147
A:Molecule type: DNA
A:Residues: 1-92 <BON>
A:Cross-references: GB:M31670; NID:9204447; PIDN:AAA41264.1; PID:9204448
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A>Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot
A:Reference number: A94090; MUID:86094338
A:Accession: B26173
A:Molecule type: mRNA
A:Residues: 1-92 <ADE>
A:Cross-references: GB:M12579; NID:9204445; PIDN:AAA41263.1; PID:9204446
R:Waller, C.C.; Marchetti, B.; Leboeuf, R.D.; Blalock, J.E.
Cell. Mol. Neurobiol. 12, 447-454, 1992
A>Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone
A:Reference number: A48410; MUID:93105480

A:Accession: A48410
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <MAT>
A:Cross-references: GB:S50870; NID:9262059; PIDN:AB24572.1; PID:9262060
A:Experimental source: thymus
A>Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBI:P:121083)
C:Genetics:
A:Introns: 47/3; 79/3
C:Function:
A:Description: stimulates pituitary secretion of luteotropin and follitropin
A>Note: gonadoliberin-associated protein may have prolactin release inhibiting activi
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-92/Product: progonaoliberin #status predicted <PGN>
F:24-33/Product: gonadoliberin #status predicted <Gln>
F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 55.8%; Score 58; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33

RESULT 8
RNAQI
gonadoliberin I - American alligator
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: A60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swan
Regul. Pept. 33, 105-116, 1991
A>Title: Primary structure of two forms of gonadotropin-releasing hormone from brains
A:Reference number: A60066; MUID:91352338
A:Accession: A60066
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 51.9%; Score 54; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.18;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
DB 2 HWSYGLRPG 10

RESULT 9
150644
gonadoliberin I precursor - chicken
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Gallus gallus (chicken)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
C:Accession: 150644; S33507
R:Dunn, I.C.; Chen, Y.; Hook, C.; Sharp, P.J.; Sang, H.M.
J. Mol. Endocrinol. 11, 19-29, 1993
A>Title: Characterization of the chicken preprogonadotropin-releasing hormone-I gene
A:Reference number: 150644; MUID:94059355
A:Accession: 150644
A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-92 <DU2>
A:Cross-references: EMBL:X69491; NID:g496326; PIDN:CAA49246.1; PID:g311612
C:Genetics: 47/3; 79/3
A:introns: 47/3; 79/3
C:Superfamily: gonadolibertin

Query Match 51.9%; Score 54; DB 2; Length 92;
Best Local Similarity 88.9%; Pred. No. 1.6;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
| | | | | | | | | |
Db 25 HWSYGLQPG 33

RESULT 10
150739
gonadotropin-releasing hormone - Cichlid (Haplochromis burtoni)

C:Species: Haplochromis burtoni
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: I50739
R:White, S.A.; Kasten, T.L.; Bond, C.T.; Adelman, J.P.; Fernald, R.D.
Proc. Natl. Acad. Sci. U.S.A. 92, 8363-8367, 1995
A:Title: Three gonadotropin-releasing hormone genes in one organism suggest novel roles
A:Reference number: I50739; MUID:95396797
A:Accession: I50739
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-98 <WHI>
A:Cross-references: EMBL:U31865; NID:g905398; PIDN:AAC59691.1; PID:g905399
C:Superfamily: gonadolibertin

Query Match 50.0%; Score 52; DB 2; Length 98;
Best Local Similarity 88.9%; Pred. No. 3.1;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
| | | | | | | | | |
Db 24 HWSYGLSPG 32

RESULT 11
A30996
orf-X protein - simian immunodeficiency virus (isolate Mne) (fragments)

N:Alternate names: VPX protein
C:Species: simian immunodeficiency virus, SIV
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Apr-1994
C:Accession: A30996
R:Henderson, L.E.; Sowder, R.C.; Copeland, T.D.; Benveniste, R.E.; Oroszian, S.
Science 241, 199-201, 1988
A:Title: Isolation and characterization of a novel protein (X-ORF product) from SIV and
A:Reference number: A30996; MUID:88264422
A:Accession: A30996
A:Molecule type: protein
A:Residues: 1-106 <HEN>
C:Genetics:
A:Gene: VPX
C:Superfamily: AIDS vpu protein

Query Match 50.0%; Score 52; DB 2; Length 106;
Best Local Similarity 60.0%; Pred. No. 3.4;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSPPP 16
| : | | | | | | | | | |
Db 88 HGAGGWRGPPPPPP 102

RESULT 12

ASLX3
vpu protein - simian immunodeficiency virus (macaque isolate)

N:Alternate names: orf-X protein
C:Species: simian immunodeficiency virus, SIV
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 16-Jul-1999
C:Accession: D28887
R:Chakrabarti, L.; Guyader, M.; Allison, M.; Daniel, M.D.; Desrosiers, R.C.; Tiollais, P.
Nature 328, 543-547, 1987
A:Title: Sequence of simian immunodeficiency virus from macaque and its relationship
A:Reference number: A28887; MUID:87287230
A:Accession: D28887
A:Molecule type: DNA
A:Residues: 1-112 <CHA>
A:Cross-references: GB:Y00277; GB:M16403; NID:g61730; PIDN:CAA68382.1; PID:g61734
C:Genetics:
A:Gene: vpu
C:Superfamily: AIDS vpu protein

C:Keywords: AIDS; immunodeficiency

Query Match 50.0%; Score 52; DB 1; Length 112;
Best Local Similarity 60.0%; Pred. No. 3.5;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSPPP 16
| : | | | | | | | | | |
Db 94 HGAGGWRGPPPPPP 108

RESULT 13

S03066
gene X protein - human T-cell lymphotropic virus type 4
C:Species: human T-cell lymphotropic virus type 4, HTLV-4
C>Date: 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999

C:Accession: S03066
R:Hahn, B.H.; Kong, L.I.; Lee, S.W.; Kumar, P.; Taylor, M.E.; Arya, S.K.; Shaw, G.M.
Nature 300, 184-186, 1987
A:Title: Relation of HTLV-4 to simian and human immunodeficiency-associated viruses.
A:Reference number: S03065
A:Accession: S03066
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-112 <HAH>
A:Cross-references: EMBL:X06391; NID:g61580; PIDN:CAA29688.1; PID:g61581
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1987
C:Superfamily: AIDS vpu protein

Query Match 50.0%; Score 52; DB 2; Length 112;
Best Local Similarity 60.0%; Pred. No. 3.5;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSSPPP 16
| : | | | | | | | | | |
Db 94 HGAGGWRGPPPPPP 108

RESULT 14

ASLST
vpu protein - simian immunodeficiency virus SIVagm (type 3, isolate STLV-3agm)

N:Alternate names: orf-X protein
C:Species: simian immunodeficiency virus SIVagm
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 31-Jan-1997
C:Accession: D26737
R:Hirsch, V.
submitted to GenBank, June 1987
A:Reference number: A26737
A:Accession: D26737
A:Molecule type: DNA
A:Residues: 1-112 <HIR>
C:Genetics:
A:Gene: vpu
C:Superfamily: AIDS vpu protein

RESULT 14

C:Keywords: AIDS; immunodeficiency

Query Match 49.0%; Score 51; DB 1; Length 112;
Best Local Similarity 60.0%; Pred. No. 4.8;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPP 16
| : | | | | | | | |
Db 94 HRAGGWRPGPPPPP 108

RESULT 15

T46289
hypothetical protein DKFZp434A10.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C:Accession: T46289

R:Duysterhoef, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000

A:Reference number: 223035

A:Accession: T46289

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-862 <AAA>

A:Cross-references: EMBL:AL137579

A:Experimental source: adult testis; clone DKFZp434A10.1

C:Genetics:

A:Note: DKFZp434A10.1

Query Match 49.0%; Score 51; DB 2; Length 862;
Best Local Similarity 80.0%; Pred. No. 36;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 7 LRPSSSPPP 16
| | | : | | | |
Db 458 LRPSSSPPP 467

Search completed: March 2, 2001, 10:55:40
Job time: 260 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:54:36 ; Search time 47.48 Seconds
(without alignments)
6.429 Million cell updates/sec

Title: US-09-306-689-3

Perfect score: 108

Sequence: 1 CPPPSEHWSYGLRPG 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	102	94.4	17 1	US-08-188-223-5
2	102	94.4	17 3	US-08-968-466-5
3	82.5	76.4	16 1	US-08-188-223-7
4	82.5	76.4	16 3	US-08-968-466-7
5	81	75.0	34 3	US-09-100-414B-82
6	81	75.0	35 3	US-09-100-414B-80
7	81	75.0	35 3	US-09-100-414B-81
8	71	65.7	49 1	US-08-387-156-4
9	71	65.7	49 2	US-08-694-865-4
10	71	65.7	49 2	US-08-878-748-4
11	71	65.7	49 2	US-09-124-491-4
12	71	65.7	544 1	US-08-387-156-10
13	71	65.7	544 2	US-08-694-865-10
14	71	65.7	544 2	US-08-878-748-10
15	71	65.7	544 2	US-09-124-491-10
16	71	65.7	699 2	US-08-694-865-16
17	71	65.7	699 3	US-09-124-491-16
18	71	65.7	977 1	US-08-387-156-8
19	71	65.7	977 2	US-08-694-865-8
20	71	65.7	977 2	US-08-878-748-8
21	71	65.7	977 2	US-09-124-491-8
22	68	63.0	14 1	US-07-690-983D-24
23	68	63.0	24 1	US-07-690-983D-43
24	68	63.0	44 1	US-07-690-983D-45
25	68	63.0	84 1	US-07-690-983D-47
26	66	61.1	90 1	US-08-341-219-19
27	66	61.1	17 1	US-07-690-983D-16
28	66	61.1	28 1	US-08-446-692-10

29	66	61.1	28 2	US-08-488-351A-10	Sequence 10, Appl
30	65	60.2	14 1	US-07-690-983D-30	Sequence 30, Appl
31	64	59.3	16 1	US-08-453-598-26	Sequence 26, Appl
32	64	59.3	16 3	US-08-521-079-26	Sequence 6, Appl
33	64	59.3	52 3	US-08-458-814-6	Sequence 6, Appl
34	64	59.3	55 3	US-08-458-814-7	Sequence 7, Appl
35	64	59.3	936 1	US-08-455-970A-12	Sequence 12, Appl
36	63	58.3	10 1	US-07-714-540-9	Sequence 9, Appl
37	63	58.3	10 1	US-07-690-983D-2	Sequence 2, Appl
38	63	58.3	10 1	US-07-690-983D-32	Sequence 32, Appl
39	63	58.3	10 1	US-08-343-883-1	Sequence 1, Appl
40	63	58.3	10 1	US-08-000-931-5	Sequence 5, Appl
41	63	58.3	10 1	US-08-428-488-22	Sequence 22, Appl
42	63	58.3	10 1	US-08-341-219-11	Sequence 11, Appl
43	63	58.3	10 1	US-08-453-588-22	Sequence 22, Appl
44	63	58.3	10 1	US-08-591-917-1	Sequence 1, Appl
45	63	58.3	10 5	5492893-1	Patent No. 5492893

ALIGNMENTS

RESULT 1
US-08-188-223-5
Sequence 5, Application US/08188223
Patent No. 5688506
GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
APPLICANT: Scibienski, Robert
TITLE OF INVENTION: Immunogens Against Gonadotropin
TITLE OF INVENTION: Releasing Hormone
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dimitrios T. Drivas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188,223
FILING DATE: 27-JAN-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8286
TELEFAX: 212-354-8113
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: YES
FEATURE:
NAME/KEY: Region
LOCATION: 1..7
OTHER INFORMATION: /note= "spacer"
FEATURE:
NAME/KEY: Region
LOCATION: 8..17
OTHER INFORMATION: /note= "immunomimic"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 17

OTHER INFORMATION: /label=glyNH2
OTHER INFORMATION: /note="glycinamide"
US-08-188-223-5

Query Match 94.4%; Score 102; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.1e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CPPPSEHWSYGLRP 16
Db 1 CPPPSEHWSYGLRP 16

RESULT 2
US-08-968-466-5
Sequence 5, Application US/08968466
Patent No. 6132720
GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
APPLICANT: Scibienski, Robert
TITLE OF INVENTION: Immunogens Against Gonadotropin
TITLE OF INVENTION: Releasing Hormone
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dimitrios T. Drivas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,466
FILING DATE: 27-JAN-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-354-8113
TELEFAX: 212-354-8113
INFORMATION FOR SEO ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: YES
FEATURE:
NAME/KEY: Region
LOCATION: 1..7
OTHER INFORMATION: /note="spacer"
FEATURE:
NAME/KEY: Region
LOCATION: 8..17
OTHER INFORMATION: /note="immunomimic"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 17
OTHER INFORMATION: /label=glyNH2
OTHER INFORMATION: /note="glycinamide"
US-08-968-466-5

Query Match 94.4%; Score 102; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.1e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CPPPSEHWSYGLRP 16
Db 1 CPPPSEHWSYGLRP 16

RESULT 3
US-08-188-223-7
Sequence 7, Application US/08188223
Patent No. 5688506
GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
APPLICANT: Scibienski, Robert
TITLE OF INVENTION: Immunogens Against Gonadotropin
TITLE OF INVENTION: Releasing Hormone
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dimitrios T. Drivas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188,223
FILING DATE: 27-JAN-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-354-8113
TELEFAX: 212-354-8113
INFORMATION FOR SEO ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: YES
FEATURE:
NAME/KEY: Region
LOCATION: 1..6
OTHER INFORMATION: /note="spacer"
FEATURE:
NAME/KEY: Region
LOCATION: 7..16
OTHER INFORMATION: /note="immunomimic"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 16
OTHER INFORMATION: /label=glyNH2
OTHER INFORMATION: /note="glycinamide"
US-08-188-223-7

Query Match 76.4%; Score 82.5; DB 1; Length 16;
Best Local Similarity 87.5%; Pred. No. 1.4e-05;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CPPPSEHWSYGLRP 16
Db 1 CPPPP-REHWSYGLRP 15

RESULT 4
US-08-968-466-7

Sequence 7, Application US/08968466
Patent No. 6132720
GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
APPLICANT: Scibienski, Robert
TITLE OF INVENTION: Immunogens Against Gonadotropin
TITLE OF INVENTION: Releasing Hormone
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dimitrios T. Drivas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,466
FILING DATE: 27-JAN-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-354-8113
TELEFAX: 212-354-8113
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
FEATURE:
NAME/KEY: Region
LOCATION: 1..6
OTHER INFORMATION: /note="spacer"
FEATURE:
NAME/KEY: Region
LOCATION: 7..16
OTHER INFORMATION: /note="immunomimic"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 16
OTHER INFORMATION: /label="GLYNH2"
OTHER INFORMATION: /note="glycinamide"
US-08-968-466-7

Query Match 76.4%; Score 82.5; DB 3; Length 16;
Best Local Similarity 87.5%; Pred. No. 1.4e-05;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CPPPSSEHWSYGLRP 16
DB 1 CPPPP-REHWSYGLRP 15

RESULT 5
US-09-100-414B-82
Sequence 82, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 80:

Query Match 75.0%; Score 81; DB 3; Length 34;
Best Local Similarity 81.2%; Pred. No. 4.8e-05;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PPPPSSEHWSYGLRPG 17
DB 19 PPXPPEHWSYGLRPG 34

RESULT 6
US-09-100-414B-80
Sequence 80, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 80:

```

; SEQUENCE CHARACTERISTICS:
;     LENGTH: 35 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: peptide
;
US-09-100-414B-80

```

Query Match	75.0%;	Score 81;	DB 3;	Length 35
Best Local Similarity	81.2%;	Pred. No. 4.9e-05;		
Matches 13; Conservative	0;	Mismatches 3;	Indels	

```
QY      2 PPPSSEHWSYGLRPG 17
        ||| | | | | | | |
Db      20 PXPXPPEHWSYGLRPG 35
```

RESULT 7
US-09-100-414B-81
; Sequence 81, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:

TITLE OF INVENTION: NOVEL LHRH PEPTIDES
 TITLE OF INVENTION: IMMUNOGENS
 NUMBER OF SEQUENCES: 106
 CORRESPONDENCE ADDRESS:
 ADDRESS: Morgan & Finnegan, L.L.P.
 STREET: 345 Park Avenue

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29, 323
REFERENCE/DOCKET NUMBER: 1151-415
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800

```

; ; SEQUENCE CHARACTERISTICS:
; ; LENGTH: 35 amino acids
; ; TYPE: amino acid
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: peptide
US-09-100-414B-81

```

Query Match	75.08;	Score 81;	DB 3;	Length 35,
Best Local Similarity	81.28;	Pred. No. 4.9e-05;		
Matches 13; Conservative	0;	Mismatches 3;	Indels	

```
QY      2 PPPSSEHWSYGLRPG 17
          ||| ||||| |||
Db      20 PXPXPPEHWSYGLRPG 35
```

RESULT 8
US-08-387-156-4
; Sequence 4, Application US/08387156
; Patent No. 5723129
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.

1 APPLICANT: REDMOND, MARK J.
 2 APPLICANT: HUGHES, HUW P.-A.
 3 TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
 4 NUMBER OF SEQUENCES: 28
 5 CORRESPONDENCE ADDRESS:
 6

STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

; APPLICATION NUMBER: US/08/387,156
 ; FILING DATE: 10-FEB-1995
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/960,932

PRIOR APPLICATION DATA: US 07/779,171
APPLICATION NUMBER:
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208

TELEPHONE: (415) 621-6393
TELEFAX: (415) 327-8231
INFORMATION FOR SEQ ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-387-156-4

Query Match	65.7%	Score 7
Best Local Similarity	80.0%	Pred. N
Matches 12; Conservative	1;	Mism

```

RESULT      9
US-08-694-865-4
; Sequence 4, Application US/08694865
; Patent No. 5837268
; GENERAL INFORMATION:

```

TITLE OF INVENTION: GNRH-LEUKOTOXIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:

STREET: 285 HAMILTON AVENUE, SUITE 100
CITY: PALO ALTO
STATE: CA

```

:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: PatentIn Release #1.0,
:
: CURRENT APPLICATION DATA:

```

```
length 49;
2; Indels 0; Gaps 0;
```

APPLICATION NUMBER: US/08/694,865
FILING DATE: 09-AUG-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-865-4

Query Match 65.7%; Score 71; DB 2; Length 49;
Best Local Similarity 80.0%; Pred. No. 0.0016;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 PPPSEHWSYGLRPG 17
1 ||:|||||
DB 22 PGSSSQHWSYGLRPG 36

RESULT 10
US-08-878-748-4
Sequence 4, Application US/08878748
Patent No. 5969126
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HOW P. A.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,748
FILING DATE: 19-JUN-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-748-4

Query Match 65.7%; Score 71; DB 2; Length 49;
Best Local Similarity 80.0%; Pred. No. 0.0016;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 PPPSEHWSYGLRPG 17
1 ||:|||||
DB 22 PGSSSQHWSYGLRPG 36

RESULT 11
US-09-124-491-4
Sequence 4, Application US/09124491
Patent No. 6022960
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANN, JOHN G.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,491
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491-4

Query Match 65.7%; Score 71; DB 3; Length 49;
Best Local Similarity 80.0%; Pred. No. 0.0016;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 PPPSEHWSYGLRPG 17
1 ||:|||||
DB 22 PGSSSQHWSYGLRPG 36

```
RESULT 12
US-08-387-156-10
; Sequence 10, Application US/08387156
; Patent No. 5723129
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HOW P.A.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/387,156
; FILING DATE: 10-FEB-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-387-156-10

Query Match          65.7%; Score 71; DB 1; Length 544;
Best Local Similarity 80.0%; Pred. No. 0.02;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      3 PPSSEHWSYGLRPG 17
      1 11:|||||
Db      515 PGSSQHSYGLRPG 529

RESULT 13
US-08-694-865-10
; Sequence 10, Application US/08694865
; Patent No. 5837268
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNIS, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
```

```
          COUNTRY: USA
          ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,865
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-694-865-10

Query Match          65.7%; Score 71; DB 2; Length 544;
Best Local Similarity 80.0%; Pred. No. 0.02;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      3 PPSSEHWSYGLRPG 17
      1 11:|||||
Db      515 PGSSQHSYGLRPG 529

RESULT 14
US-08-878-748-10
; Sequence 10, Application US/08878748
; Patent No. 5969126
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HOW P.A.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,748
; FILING DATE: 19-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
```

REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-748-10

Query Match 65.7%; Score 71; DB 2; Length 544;
Best Local Similarity 80.0%; Pred. No. 0.02;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 PPPSEHWSYGLRPG 17
| ||:|||||||
DB 515 PGSSQHSYGLRPG 529

RESULT 15
US-09-124-491-10
Sequence 10, Application US/09124491
Patent No. 6022960
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANNIS, JOHN G.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,491
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491-10

Query Match 65.7%; Score 71; DB 3; Length 544;
Best Local Similarity 80.0%; Pred. No. 0.02;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 PPPSEHWSYGLRPG 17
| ||:|||||||
DB 515 PGSSQHSYGLRPG 529

Search completed: March 2, 2001, 10:54:36
Job time: 601 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 11:00:45 ; Search time 97.13 Seconds
(without alignments)
20.514 Million cell updates/sec

Title: US-09-306-689-3
Perfect score: 108
Sequence: 1 CPPPSEHMSYGLRPG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_15:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.todent:*
13: sp.virus:*
14: sp.vertebrate:*
15: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	55.6	91	13	09PRH0
2	54	50.0	87	13	09YI26
3	54	50.0	95	13	073812
4	54	50.0	99	13	091A10
5	54	50.0	114	6	097655
6	52	48.1	90	13	091A02
7	52	48.1	110	6	097686
8	52	48.1	148	4	09WZ81
9	52	48.1	2357	5	09UIM8
10	51	47.2	33	13	09W760
11	51	47.2	33	13	09PT34
12	51	47.2	82	13	092094
13	51	47.2	82	13	09W761
14	51	47.2	82	13	091800
15	51	47.2	82	13	0918P9
16	51	47.2	88	13	09PSY9
17	51	47.2	90	13	091A09
18	51	47.2	107	6	09YI26
19	51	47.2	327	4	09UHN2

20	51	47.2	327	4	09UBV7	09ubv7 homo sapien
21	50.5	46.8	2942	12	09WB76	09wb76 gp virus c
22	50	46.3	86	13	042471	042471 carassius a
23	50	46.3	86	13	09PT25	09pt25 oncorhynch
24	50	46.3	230	5	09W3V7	09w3v7 dirosophila
25	50	46.3	497	5	018662	018662 caenorhabd1
26	49.5	45.8	89	2	09ZNI3	09zni3 pseudomonas
27	49.5	45.8	437	6	09RS25	09rs25 bos taurus
28	49.5	45.8	473	5	09YOS9	09yos9 dirosophila
29	49	45.4	816	5	097214	097214 leishmania
30	49	45.4	880	5	09W575	09w575 dirosophila
31	48	44.4	348	5	09VSB8	09vsb8 dirosophila
32	47.5	44.0	305	5	015844	015844 leishmania
33	47.5	44.0	1044	4	09P206	09p206 homo sapien
34	47	43.5	75	10	001948	001948 lycopersico
35	47	43.5	190	2	086520	086520 streptomyce
36	47	43.5	226	2	069910	069910 streptomyce
37	47	43.5	322	10	009084	009084 lycopersico
38	47	43.5	325	10	081467	081467 arabidopsis
39	47	43.5	327	5	09YTP2	09ytp2 caenorhabd1
40	47	43.5	612	4	075802	075802 homo sapien
41	47	43.5	788	5	096416	096416 dirosophila
42	46.5	43.1	267	4	013261	013261 homo sapien
43	46	42.6	164	2	09PCB4	09pcb4 xyliella fas
44	46	42.6	201	10	09XIN6	09xin6 arabidopsis
45	46	42.6	252	10	09SD40	09sd40 arabidopsis

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	91 AA
09PRH0	09PRH0			
AC	09PRH0			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)			
DE	PREPRO-MGNRH PRECURSOR.			
OS	Anguilla japonica (Japanese eel).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;			
OC	NCBI_TaxID=7937;			
OX	[1]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=BRAIN;			
RA	Okubo K., Suetake H., Aida K.;			
RT	"Expression of two gonadotropin-releasing hormone (GNRH) precursor			
RT	genes in various tissues of the Japanese eel and evolution of GNRH.";			
RL	Zool. Sci. 16:471-478(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Okubo K., Suetake H., Aida K.;			
RT	"A splicing variant for the prepro-mammalian gonadotropin-releasing			
RT	hormone (prepro-mGNRH) mRNA is present in the brain and various			
RT	peripheral tissues of the Japanese eel.";			
RL	Zool. Sci. 16:645-651(1999).			
DR	EMBL; AB026989; BAA82608.1; -			
DR	EMBL; AB026991; BAA83597.1; -			
DR	INTERPRO: IPR02012; -			
DR	PFAM; PF00446; GNRH; 1.			
DR	PROSITE; PS00473; GNRH; 1.			
KW	SIGNAL.			
FT	1	22		POTENTIAL.
FT	CHAIN	23	32	MGNRH.
FT	CHAIN	33	91	GNRH ASSOCIATED PEPTIDE.
FT	SEQUENCE	91 AA;	9893 MW;	BA15C9DC08434A7B CRC64;

Query Match 55.6%; Score 60; DB 13; Length 91;
Best local similarity 90.0%; Pred. No. 0.11;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 EHWSTGLRPG 17
:|||||
Db 23 QHWSTGLSPG 32

RESULT 2
09Y126 PRELIMINARY; PRT; 87 AA.

AC 09Y126:
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN)
DE (FRAGMENT).
OS Sparus aurata (Gilthead sea bream).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Sparus.
OX NCBI_TaxID=8175;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Ovary;
RA Nabissi M.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
FOLLICLE-STIMULATING HORMONES.
DR EMBL; AF046801; AAD02427.1; -.
DR INTERPRO; IPR002012; -.
DR PFAM; PF00446; GNRH; 1.
RW PROSITE; PS00473; GNRH; 1.
KW Hormone; Amidation.
FT NON_TER 1 1
FT SEQUENCE 87 AA; 9871 MW; 0D246353D96782A CRC64;

Query Match 50.0%; Score 54; DB 13; Length 87;
Best Local Similarity 80.0%; Pred. NO. 0.75;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 EHWSTGLRPG 17
:|||||
Db 21 QHWSTGLSPG 30

RESULT 3
073812 PRELIMINARY; PRT; 95 AA.

AC 073812:
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN).
OS Morone saxatilis (Striped bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Moronidae; Morone.
OX NCBI_TaxID=34816;
RN [1]
RP SEQUENCE FROM N.A.
RC Chow M.M., Kight K.E., Gotthelf Y., Alok D., Zohar Y.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
FOLLICLE-STIMULATING HORMONES.
DR EMBL; AF056314; AAD03817.1; -.
DR INTERPRO; IPR002012; -.
DR PFAM; PF00446; GNRH; 1.
RW PROSITE; PS00473; GNRH; 1.

DR PRODOM; PD005581; -; 1.
KW Hormone; Amidation.
SQ SEQUENCE 95 AA; 10411 MW; 980C6988FC279BFC CRC64;

Query Match 50.0%; Score 54; DB 13; Length 95;
Best Local Similarity 80.0%; Pred. NO. 0.81;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 EHWSTGLRPG 17
:|||||
Db 23 QHWSTGLSPG 32

RESULT 4
09IA10 PRELIMINARY; PRT; 99 AA.

AC 09IA10:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GONADOTROPIN-RELEASING HORMONE SEABREAM ISOFORM.
OS Dicentrarchus labrax (European sea bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Moronidae; Dicentrarchus.
OX NCBI_TaxID=13489;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Gonzalez-Martinez D., Madigou T., Zmora N., Anglade I., Zanuy S.,
RA Zohar Y., Elizur A., Munoz-Cueto J.A., Kah O.;
RW "differential expression of three different prepro-GNRH
(Gonadotrophin-releasing hormone) messengers in the brain of the
European sea bass (Dicentrarchus labrax).";
RL Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Zmora N., Zohar Y., Elizur A.;
RW "3 GNRH form in the seabass Dicentrarchus labrax.";
RL Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF224279; AAF62898.1; -.
SQ SEQUENCE 99 AA; 10758 MW; EC8AE9C93CC02904 CRC64;

Query Match 50.0%; Score 54; DB 13; Length 99;
Best Local Similarity 80.0%; Pred. NO. 0.85;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 EHWSTGLRPG 17
:|||||
Db 27 QHWSTGLSPG 36

RESULT 5
097655 PRELIMINARY; PRT; 114 AA.

AC 097655:
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN)
DE PRECURSOR.
GN GNRH2.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.

RA White R.B., Urbanski H.F., Fernald R.D.;
 RT "A second gene for gonadotropin-releasing hormone is expressed in the
 RL rhesus macaque (Abstract #632.18)".
 CC Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC FOLLICLE-STIMULATING HORMONES.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 DR EMBL: AF097356; AAD09106.1; -.
 DR INTERPRO: IPR002012; -.
 DR PIRAM: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 DR PRODOM: PD029787; -; 1.
 KW Hormone; Amidation.
 SQ SEQUENCE 114 AA; 12533 MW; 8B70D690D5BD5103 CRC64;

Query Match 50.0%; Score 54; DB 6; Length 114;
 Best Local Similarity 50.0%; Pred. No. 0.97;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 PPPPSSEHWSYGLRPG 17
 Db 19 PGPEAQMWSHGWPY 34

RESULT 6
 O91AU2 PRELIMINARY; PRT; 90 AA.
 AC O91AU2;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GONADOTROPIN-RELEASING HORMONE.
 OS Rana dybowskii (Frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
 OX NCBI_TaxID=71582;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RA Yoo M.S., Kang H.M., Choi H.S., Chun S.Y., Troskie B., Millar R.P.,
 RA Kwon H.B.;
 RT "Molecular Cloning, Distribution and Pharmacological Characterization
 RT of a Novel Gonadotropin-releasing Hormone (Trp8)GnRH in Frog Brain";
 RL Mol. Cell. Endocrinol. 0:0-0(2000).
 DR EMBL: AF139111; AAF44343.1; -.
 SQ SEQUENCE 90 AA; 10368 MW; C3D573E78B52ABFA CRC64;

Query Match 48.1%; Score 52; DB 13; Length 90;
 Best Local Similarity 80.0%; Pred. No. 1.5;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 8 EHWSYGLRPG 17
 Db 25 QHWSYGLMPG 34

RESULT 7
 O97686 PRELIMINARY; PRT; 110 AA.
 AC O97686;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN).
 GN GNRH2.
 OS Suncus murinus (House shrew).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Insectivora; Soricidae; Suncus.
 OX NCBI_TaxID=9378;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=BRAIN.
 RA White R.B., Kastren T.L., White S.A., Rissman E.F., Fernald R.D.;
 RT "GnRH-II cDNA expression in the musk shrew";
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC FOLLICLE-STIMULATING HORMONES.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 DR EMBL: AF107315; AAD09114.1; -.
 DR INTERPRO: IPR002012; -.
 DR PIRAM: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 DR PRODOM: PD029787; -; 1.
 KW Hormone; Amidation.
 SQ SEQUENCE 110 AA; 12120 MW; A8986905FB83D9DB CRC64;

Query Match 48.1%; Score 52; DB 6; Length 110;
 Best Local Similarity 50.0%; Pred. No. 1.8;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 PPPPSSEHWSYGLRPG 17
 Db 21 PGPLKAQWSHGWPY 36

RESULT 8
 O9N281 PRELIMINARY; PRT; 148 AA.
 AC O9N281;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE UNCHARACTERIZED BONE MARROW PROTEIN BM041.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BONE MARROW;
 RA Zhao M., Gu J., Li N., Peng Y., Han Z., Chen Z.;
 RT "A novel gene expressed in human bone marrow";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF217517; AAF67628.1; -.
 SQ SEQUENCE 148 AA; 15385 MW; 56BFE6A15935A2E CRC64;

Query Match 48.1%; Score 52; DB 4; Length 148;
 Best Local Similarity 47.1%; Pred. No. 2.4;
 Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 CPPPSSEHWSYGLRPG 17
 Db 41 CPPPGAPGNGNPAFP 57

RESULT 9
 O901M8 PRELIMINARY; PRT; 2357 AA.
 AC O901M8;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE CLASS VII UNCONVENTIONAL MYOSIN.
 GN MYO1.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Dictyostellida; Dictyostellum.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE OF 254-349 FROM N.A.
 RX MEDLINE=95023928; PubMed=7937787;
 RA Titus M.A., Kuspa A., Loomis W.F.;
 RT "Discovery of myosin genes by physical mapping in Dictyostelium";

RL Proc. Natl. Acad. Sci. U.S.A. 91:9446-9450(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Titus M.A.;
 RT "A class VII unconventional myosin is required for phagocytosis."
 RL Curr. Biol. 9:1297-1303(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Titus M.A., Loomis W.F.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Titus M.A.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: L35321; AAF06035.1; -
 DR HSSP: P08799; 1MND.
 DR INTERPRO: IPR000048; -
 DR INTERPRO: IPR000159; -
 DR INTERPRO: IPR000299; -
 DR INTERPRO: IPR000857; -
 DR INTERPRO: IPR001452; -
 DR INTERPRO: IPR001609; -
 DR PFAM: PF000018; SH3; 1.
 DR PFAM: PF00063; myosin_head; 1.
 DR PFAM: PF00612; IQ; 1.
 DR PFAM: PF00784; MYTH; 2.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PROSITE: PS50002; SH3; 1.
 DR PROSITE: PS50057; BAND; 41.3; 2.
 SO SEQUENCE 2357 AA; 268281 MW; FC667F12F207B0C2 CRC64;

Query Match 48.1%; Score 52; DB 5; Length 2357;
 Best Local Similarity 60.0%; Pred. No. 34;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 PPPSSSEHMYGLRP 16
 Db 1027 PPPSSSSMDGLRP 1041

RESULT 10
 Q9W7G0 PRELIMINARY; PRT; 33 AA.
 AC Q9W7G0;
 DT 01-NOV-1999 (TREMBLREL. 12, Created)
 DT 01-NOV-1999 (TREMBLREL. 12, Last sequence update)
 DT 01-MAY-2000 (TREMBLREL. 13, Last annotation update)
 DE GONADOTROPIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN)
 DE (FRAGMENT).
 GN GNRH2.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OC NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99312119; PubMed-10385393;
 RA Von Schalburg K.R.; Sherwood N.M.;
 RT "Regulation and expression of gonadotropin-releasing hormone gene
 differs in brain and gonads in rainbow trout."
 RL Endocrinology 140:3012-3024(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Von Schalburg K.R.; Sherwood N.M.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC FOLLICLE-STIMULATING HORMONES.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 DR EMBL: AF110993; AAD3463.1; -
 DR INTERPRO: IPR002012; -
 DR PFAM: PF00446; GNRH; 1.

DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation.
 FT NON_TER 33
 SO SEQUENCE 33 AA; 3668 MW; 099C825E4A72A3BB CRC64;

Query Match 47.2%; Score 51; DB 13; Length 33;
 Best Local Similarity 72.7%; Pred. No. 0.8;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 SEHWSYGLRPG 17
 Db 23 SQHWSYGLWLP 33

RESULT 11
 Q9PT34 PRELIMINARY; PRT; 33 AA.
 AC Q9PT34;
 DT 01-MAY-2000 (TREMBLREL. 13, Created)
 DT 01-MAY-2000 (TREMBLREL. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLREL. 14, Last annotation update)
 DE GONADOTROPIN-RELEASING HORMONE (FRAGMENT).
 GN GNRH1.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OC NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99312119; PubMed-10385393;
 RA Von Schalburg K.R.; Sherwood N.M.;
 RT "Regulation and expression of gonadotropin-releasing hormone gene
 differs in brain and gonads in rainbow trout."
 RL Endocrinology 140:3012-3024(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Von Schalburg K.R.; Sherwood N.M.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF110533; AAD3461.1; -
 DR INTERPRO: IPR002012; -
 DR INTERPRO: IPR002047; -
 DR PFAM: PF00446; GNRH; 1.
 DR PROSITE: PS00256; AKH; UNKNOWN_1.
 DR PROSITE: PS00473; GNRH; 1.
 FT NON_TER 33
 SO SEQUENCE 33 AA; 3741 MW; 1FE1535E742B7EBB CRC64;

Query Match 47.2%; Score 51; DB 13; Length 33;
 Best Local Similarity 72.7%; Pred. No. 0.8;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 SEHWSYGLRPG 17
 Db 23 SQHWSYGLWLP 33

RESULT 12
 Q92094 PRELIMINARY; PRT; 82 AA.
 AC Q92094;
 DT 01-NOV-1996 (TREMBLREL. 01, Created)
 DT 01-NOV-1996 (TREMBLREL. 01, Last sequence update)
 DT 01-OCT-2000 (TREMBLREL. 15, Last annotation update)
 DE GONADOTROPIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN)
 DE PRECURSOR.
 GN PREPRO-GNRH-1.
 OS Oncorhynchus nerka (Sockeye salmon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

OX NCBI_TaxID=8023;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=NIKKO; TISSUE=BRAIN;
 RA MEDLINE=96020547; PubMed=8546809;
 RX Ashihara M., Suzuki M., Kubokawa K., Yoshitura Y., Kobayashi M.,
 RU Urano A., Aida K.;
 RT "Two differing precursor genes for the salmon-type gonadotropin-
 releasing hormone exist in salmonids.";
 RL J. Mol. Endocrinol. 15:1-9(1995).
 CC -I- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC FOLLICLE-STIMULATING HORMONES.
 DR EMBL: D31868; BAA06666.1; -
 DR INTERPRO: IPR002047; -
 DR INTERPRO: IPR002047; -
 DR PFAM: PF00446; GNRH: 1.
 DR PROSITE: PS00256; AKH; UNKNOWN_1.
 DR PRODOM: PD006549; -; 1.
 DR SIGNAL: Hormone; Amidation.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 33 GNRH-ASSOCIATED PEPTIDE.
 FT CHAIN 37 82 GNRH-ASSOCIATED PEPTIDE.
 SQ SEQUENCE 82 AA; 9126 MW; C64044EA521B2B8B CRC64;

Query Match 47.2%; Score 51; DB 13; Length 82;
 Best Local Similarity 72.7%; Pred. No. 1.9;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17
 Db 23 SQHWSYGLPG 33

RESULT 13
 ID 09W7G1 PRELIMINARY; PRT: 82 AA.
 AC 09W7G1;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN).
 GN GNRH1.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99312119; PubMed=10385393;
 RX Von Schalburg K.R., Sherwood N.M.;
 RT "Regulation and expression of gonadotropin-releasing hormone gene
 RT differs in brain and gonads in rainbow trout.";
 RL Endocrinology 140:3012-3024(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Von Schalburg K.R., Sherwood N.M.;
 RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
 CC -I- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC FOLLICLE-STIMULATING HORMONES.
 DR EMBL: AF110992; AADA3462.1; -
 DR INTERPRO: IPR002012; -
 DR INTERPRO: IPR002047; -
 DR PFAM: PF00446; GNRH: 1.
 DR PROSITE: PS00256; AKH; UNKNOWN_1.
 DR PROSITE: PS00473; GNRH: 1.
 DR Hormone; Amidation.
 SQ SEQUENCE 82 AA; 9232 MW; 7595B4FCC65FDFD6 CRC64;

Query Match 47.2%; Score 51; DB 13; Length 82;
 Best Local Similarity 72.7%; Pred. No. 1.9;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17
 Db 23 SQHWSYGLPG 33

RESULT 14
 ID 091800 PRELIMINARY; PRT: 82 AA.
 AC 091800;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE PRO-SGNRH-1.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RA Ferriere F., Bailhache T., Jégo P.;
 RT "Oncorhynchus mykiss gnrh-1 cDNA from brain.";
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF232212; AAF91280.1; -
 SQ SEQUENCE 82 AA; 9198 MW; 7595A0B896556A69 CRC64;

Query Match 47.2%; Score 51; DB 13; Length 82;
 Best Local Similarity 72.7%; Pred. No. 1.9;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17
 Db 23 SQHWSYGLPG 33

RESULT 15
 ID 0918P9 PRELIMINARY; PRT: 82 AA.
 AC 0918P9;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE PRO-SGNRH-1.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RA Ferriere F., Bailhache T., Jégo P.;
 RT "Oncorhynchus mykiss gnrh-1 cDNA in the brain.";
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF232213; AAF91281.1; -
 SQ SEQUENCE 82 AA; 9203 MW; 8053F4F221A0F08 CRC64;

Query Match 47.2%; Score 51; DB 13; Length 82;
 Best Local Similarity 72.7%; Pred. No. 1.9;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17
 Db 23 SQHWSYGLPG 33

Fri Mar 2 12:19:30 2001

us-09-306-689-3.rspt

Page 6

Search completed: March 2, 2001, 11:00:47
Job time: 457 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 11:01:19 ; Search time 32.03 seconds
(without alignments)
17.140 Million cell updates/sec

Title: US-09-306-689-3
Perfect score: 108
Sequence: 1 CPPPSEHWSYGLRPG 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues
Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Swissprot_39:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	63.9	92	1	GON1_HUMAN
2	68	63.0	67	1	GON1_MACMU
3	68	63.0	90	1	GON1_MOUSE
4	68	63.0	91	1	GON1_PIG
5	68	63.0	92	1	GON1_RAT
6	61	56.5	89	1	GON1_XENLA
7	61	56.5	92	1	GON1_TUPGB
8	60	55.6	61	1	GON1_SHEEP
9	60	55.6	63	1	GON1_MESAU
10	57	52.8	92	1	GON1_CAVPO
11	57	52.8	92	1	GON1_CHICK
12	56	51.9	10	1	GON1_ALMT
13	56	51.9	94	1	GON1_HAPBU
14	55	50.9	80	1	GON1_CLAGA
15	54	50.0	95	1	GON1_PAGMA
16	54	50.0	95	1	GON1_SPRAU
17	54	50.0	114	1	GON2_TUPGB
18	51	47.2	74	1	GON3_ONCMT
19	51	47.2	74	1	GON3_ONCMT
20	51	47.2	82	1	GON3_ONCMT
21	51	47.2	82	1	GON3_ONCMT
22	51	47.2	82	1	GON3_ONCMT
23	51	47.2	89	1	GON3_PORNO
24	51	47.2	90	1	GON3_HAPBU
25	51	47.2	90	1	GON3_PAGMA
26	51	47.2	90	1	GON3_SPRAU
27	50	46.3	86	1	GON2_CARAU
28	50	46.3	86	1	GON2_ONCMT
29	50	46.3	86	1	GON2_RUTRU
30	49.5	45.8	437	1	PTB2_BOVIN
31	49	45.4	120	1	GON2_HUMAN
32	48	44.4	906	1	KBF2_CHICK
33	47	43.5	10	1	GON3_ONCMT

ALIGNMENTS

34	47	43.5	94	1	GON3_CARAU	P51917 carcassius a
35	47	43.5	94	1	GON3_RUTRU	O92106 rutillus rut
36	47	43.5	396	1	FLO_ANTMA	P23915 antirrhinum
37	47	43.5	757	1	CIFE_HUMAN	Q14003 homo sapien
38	47	43.5	769	1	CIFE_MOUSE	O63959 mus musculu
39	47	43.5	889	1	CIFE_RAT	O01956 rattus norv
40	46.5	43.1	250	1	CPBL_CHICK	P81475 gallus gall
41	46	42.6	86	1	GON2_CLAGA	P43306 claritas gar
42	46	42.6	95	1	DAMS_HUMAN	O99013 homo sapien
43	46	42.6	551	1	CAP_SCHPO	P36621 schizosach
44	46	42.6	756	1	PIP6_RAT	P10688 rattus norv
45	45	41.7	383	1	DEMA_HUMAN	O08495 homo sapien

RESULT 1
GON1_HUMAN STANDARD: PRT: 92 AA.
AC P01148;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
(LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
HORMONE I) (GNRH I) (LULIBERIN I) (GNADORELIN); GNRH-ASSOCIATED
PEPTIDE I].
GN GNRH1 OR GNRH OR LHRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-8936682; PubMed-2671939;
RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT "The complete nucleotide sequence of the human gonadotropin-releasing
hormone gene."
RL Nucleic Acids Res. 17:6403-6403(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-8609438; PubMed-2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Characterization of cDNA for precursor of human luteinizing hormone
releasing hormone."
RL Nature 311:666-668(1984).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-85012739; PubMed-6090951;
RA Seeburg P.H., Adelman J.P.;
RT "Characterization of cDNA for precursor of human luteinizing hormone
releasing hormone."
RL Nature 311:666-668(1984).
RN [4]
RP SEQUENCE OF 24-33.
RX MEDLINE-83126573; PubMed-6760865;
RA Tan L., Rousseau P.;
RT "The chemical identity of the immunoreactive LHRH-like peptide
biosynthesized in the human placenta."
RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
CC - FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
HORMONES.
CC - PHARMACEUTICAL: AVAILABLE UNDER THE NAMES FACTREL (AVERST LABS),
LUTREPISE OR LUTRELEF (FERRING PHARMACEUTICALS) AND RELISORM
(SEERON).
CC - SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: X01059; CAA25526.1; -
 DR EMBL: M12578; AAA35916.1; -
 DR EMBL: X15215; CAA33285.1; -
 DR PIR: A01410; RHUG.
 DR PIR: A26173; A26173.
 DR PIR: S05308; S05308.
 DR MIM: 152760; -
 DR INTERPRO: IPR002012; -
 DR PFAM: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KM Placenta; Pharmaceutical; Signal.

FT SIGNAL 1 23
 FT CHAIN 24 92 PROGNADOLIBERIN I.
 FT PEPTIDE 24 33 GONADOLIBERIN I.
 FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 ACTIVITY.

FT MOD_RES 24 24 PYROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
 FT CONFLICT 16 16 W -> S (IN REF. 3).
 SQ SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;

Query Match 63.9%; Score 69; DB 1; Length 92;
 Best Local Similarity 70.6%; Pred. No. 0.0054;
 Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 CPPPSEHWSYGLRPG 17
 DB 17 CVEGCSGHSYGLRPG 33

RESULT 2
 GONL_MACMU STANDARD; PRT; 67 AA.
 ID GONL_MACMU
 AC P55247;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
 DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
 DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]
 DE (FRAGMENT).
 DE GNRH1 OR GNRH OR LHRH.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheinae; Macaca.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-HYPOTHALAMUS;
 RX MEDLINE-95124501; Pubmed-7545971;
 RA Ma Y.J., Costa M.E., Ojeda S.R.;
 RT "developmental expression of the genes encoding transforming growth
 RT factor alpha and its receptor in the hypothalamus of female rhesus
 RT macaques.";
 RL Neuroendocrinology 60:346-359(1994).
 CC CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

 DR EMBL: S75918; AAB33096.1; -
 DR INTERPRO: IPR002012; -
 DR PFAM: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KM Signal.

FT SIGNAL 1 1
 FT NON_TER 67 67
 FT CHAIN 6 5 BY SIMILARITY.
 FT PEPTIDE 6 15 PROGNADOLIBERIN I.
 FT PEPTIDE 19 67 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 ACTIVITY (BY SIMILARITY).
 FT MOD_RES 6 6 PYROLIDONE CARBOXYLIC ACID (BY
 SIMILARITY).
 FT MOD_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
 SIMILARITY).
 FT NON_TER 67 67
 SQ SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;

Query Match 63.0%; Score 68; DB 1; Length 67;
 Best Local Similarity 91.7%; Pred. No. 0.0054;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 SSEHWSYGLRPG 17
 DB 4 SSQHSYGLRPG 15

RESULT 3
 GONL_MOUSE STANDARD; PRT; 90 AA.
 ID GONL_MOUSE
 AC P13562;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
 DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
 DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR
 DE I].
 DE GNRH1 OR GNRH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-87069928; Pubmed-3024317;
 RA Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,
 RA Phillips H.S., Nikolic K., Seeburg P.H.;
 RT "A deletion truncating the gonadotropin-releasing hormone gene is
 RT responsible for hypogonadism in the hpg mouse.";
 RL Science 234:1366-1371(1986).
 CC CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: M14872; AAA37717.1; -
 DR MGD: MGI:95789; GNRH.
 DR INTERPRO: IPR002012; -
 DR PFAM: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.

KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 90
 FT PEPTIDE 22 31
 FT PEPTIDE 35 90
 FT ACT_SITE 24 24
 FT MOD_RES 22 22
 FT MOD_RES 31 31
 SQ SEQUENCE 90 AA; 10337 MW; 1C0766FA826E4D9 CRC64;
 Query Match 63.0%; Score 68; DB 1; Length 90;
 Best Local Similarity 91.7%; Pred. No. 0.0072; Mismatches 0; Indels 0; Gaps 0;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 6 SSEARCH: 17
 DB 20 SSEARCH: 31
 RESULT 4
 GONL_PIG STANDARD; PRT; 91 AA.
 AC P49921:
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROCONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
 DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
 DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
 GN GNRH1 OR GNRH.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-HYPOTHALAMUS;
 RA Weesner G.D., Matteri R.L., Becker B.A.;
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 24-33.
 RX MEDLINE-72114303; PubMed-4946067;
 RA Baba Y., Matsuo H., Schally A.V.;
 RT "Structure of the porcine LH- and FSH-releasing hormone. II.
 RT Confirmation of the proposed structure by conventional sequential
 RT analyses.";
 RL Biochem. Biophys. Res. Commun. 44:459-463(1971).
 RN [3]
 RP SYNTHESIS OF GONADOLIBERIN.
 RX MEDLINE-72065376; PubMed-4942726;
 RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
 RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
 RT phase method.";
 RL Biochem. Biophys. Res. Commun. 45:822-827(1971).
 RN [4]
 RP SYNTHESIS OF GONADOLIBERIN.
 RX MEDLINE-72117544; PubMed-4946275;
 RA Baba Y., Arimura A., Schally A.V.;
 RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";
 RL Biochem. Biophys. Res. Commun. 45:483-487(1971).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: L32864; AAA1066.1; -
 DR PIR: A01411; RHPG.
 DR INTERPRO: IPR002012; -
 DR PFAM: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 KW Placenta; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 91
 FT PEPTIDE 24 33
 FT PEPTIDE 34 91
 FT ACT_SITE 26 26
 FT MOD_RES 24 24
 FT MOD_RES 33 33
 SQ SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;
 Query Match 63.0%; Score 68; DB 1; Length 91;
 Best Local Similarity 91.7%; Pred. No. 0.0073; Mismatches 0; Indels 0; Gaps 0;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 6 SSEARCH: 17
 DB 22 SSEARCH: 33
 RESULT 5
 GONL_RAT STANDARD; PRT; 92 AA.
 AC P07490;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PROCONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
 DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
 DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR
 DE I].
 GN GNRH1 OR GNRH.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-86094338; PubMed-2867548;
 RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
 RT "Isolation of the gene and hypothalamic cDNA for the common precursor
 RT of gonadotropin-releasing hormone and prolactin release-inhibiting
 RT factor in human and rat.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89384651; PubMed-2476669;
 RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
 RT "The rat gonadotropin-releasing hormone: SH locus: structure and
 RT hypothalamic expression.";
 RL Mol. Endocrinol. 3:1257-1262(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE-RHYMUS.
 RX MEDLINE-93105480; PubMed-1468115;
 RA Maier C.C., Marchetti B., Lebeouf R.D., Blalock J.E.;
 RT "Thymocytes express a mRNA that is identical to hypothalamic
 RT luteinizing hormone-releasing hormone mRNA.";
 RL Cell. Mol. Neurobiol. 12:447-454(1992).
 RN [4]
 RP SEQUENCE OF 1-47 FROM N.A.
 RC TISSUE=HEART;
 RX MEDLINE-87149087; PubMed-3547652;
 RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
 RT "Two mammalian genes transcribed from opposite strands of the same

```

RT DNA locus.
RL Science 235:1514-1517(1987).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC CC SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S50870; AAB24572.1; -.
DR EMBL; M12579; AAA41263.1; -.
DR EMBL; M31670; AAA41264.1; -.
DR EMBL; M15527; AAA42141.1; ALT_SEQ.
DR EMBL; M15529; AAA42139.1; -.
DR EMBL; M15528; -. NOT_ANNOTATED_CDS.
DR PIR; B26173; RHRTG.
DR PIR; A48410; A48410.
DR INTERPRO; IPR002012; -.
DR PFAM; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
DR KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 1 23
FT PEPTIDE 24 92 PROGONADOLIBERIN I.
FT PEPTIDE 37 92 GONADOLIBERIN I.
FT ACT_SITE 26 26 PROLACTIN RELEASE-INHIBITING FACTOR I.
FT APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;

Query Match 63.0%; Score 68; DB 1; Length 92;
Best Local Similarity 91.7%; Pred. No. 0.0074;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 SEHMSYGLRPG 17
   ||:|||||
Db 22 SSOHMSYGLRPG 33

RESULT 6
GONL_XENLA STANDARD; PRT; 89 AA.
AC P45656;
ID GONL_XENLA STANDARD; PRT; 89 AA.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE (LH-RH) (LULIBERIN I).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-FORERAIN;
RX MEDLINE-94185563; PubMed-8137750;
RA Hayes W.P., Wasy S., Batley J.F.;
RT "The frog gonadotropin-releasing hormone-I (GNRH-I) gene has a
RT mammalian-like expression pattern and conserved domains in
RT GNRH-associated peptide, but brain onset is delayed until
RT metamorphosis."
RT Endocrinology 134:1835-1844(1994).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.

```

```

CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC CC SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L28040; AAA49728.1; -.
DR INTERPRO; IPR002012; -.
DR PFAM; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
DR KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 23
FT CHAIN 1 23
FT PEPTIDE 24 89 PROGONADOLIBERIN I.
FT PEPTIDE 37 89 GONADOLIBERIN I.
FT CHAIN 37 89 GONADOTROPIN-RELEASING HORMONE ASSOCIATED
FT PEPTIDE 37 85 GNRH-ASSOCIATED PEPTIDE I (GAP).
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE 89 AA; 10246 MW; 6F4F36FBAED04284 CRC64;

Query Match 56.5%; Score 61; DB 1; Length 89;
Best Local Similarity 81.8%; Pred. No. 0.061;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 SEHMSYGLRPG 17
   :|||||
Db 23 AQHMSYGLRPG 33

RESULT 7
GONL_TUPGB STANDARD; PRT; 92 AA.
AC Q95335;
ID GONL_TUPGB STANDARD; PRT; 92 AA.
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LH-RH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
GN GNRH1 OR GNRH.
OS Tupia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupai.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-HYPOTHALAMUS;
RX MEDLINE-97079639; PubMed-8921350;
RA Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGNRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GNRH gene expression in a
RT placental mammal."
RT Gen. Comp. Endocrinol. 104:7-19(1996).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC CC SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```



```

DR EMBL: U63326; AAB16837.1;
DR INTERPRO: IPR002012;
DR PFAM: PF00446; GNRH: 1.
DR PROSITE: PS00473; GNRH: 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23 BY SIMILARITY.
FT CHAIN 24 92 PROGNADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
FT SEQUENCE 92 AA; 10197 MW; 4FDF2C58CF5F63B CRC64;
SQ
Query Match 56.5%; Score 61; DB 1; Length 92;
Best Local Similarity 81.8%; Pred. No. 0.063;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 7 SEHMSYGLRPG 17
Db 23 AQHMSYGLRPG 33
RESULT 8
GONL_SHEEP STANDARD; PRT; 61 AA.
ID GONL_SHEEP
AC 028588;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]
DE (FRAGMENT)
GN GNRH1 OR GNRH OR LHRH.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
RN [1]
RP SEQUENCE OF 12-61 FROM N.A.
RC STRAIN-WESTERN RANGE; TISSUE-HYPOTHALAMUS;
RA Rodriguez R.E.; Wise M.E.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-10.
RA MEDLINE:72094314; PubMed=4550508;
RA Burrows R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,
RA Fellows R., Blackwell R., Vale W., Guillemin R.;
RA Primary structure of the ovine hypothalamic luteinizing hormone-
RA releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass
RT spectrometry-decapeptide-Edman degradation)";
RT Proc. Natl. Acad. Sci. U.S.A. 69:282-282(1972).
RL
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 002517; AAA03433.1;
CC PIR: A93780; RSHSG.

```

```

DR INTERPRO: IPR002012;
DR PFAM: PF00446; GNRH: 1.
DR PROSITE: PS00473; GNRH: 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
FT NON_TER 1 1
FT CHAIN 1 >61 PROGNADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >61 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).
FT NON_TER 61 61
FT SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;
SQ
Query Match 55.6%; Score 60; DB 1; Length 61;
Best Local Similarity 90.0%; Pred. No. 0.057;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 8 EHSYGLRPG 17
Db 1 QHMSYGLRPG 10
RESULT 9
GONL_MESAU STANDARD; PRT; 63 AA.
ID GONL_MESAU
AC 009163;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]
DE (FRAGMENT)
GN GNRH1 OR GNRH OR LHRH.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
RN [1]
RP SEQUENCE FROM N.A.
RA Jansen H.T., Stevens P.J., Zeltner P., Lehman M.N.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U91938; AAB51302.1;
CC INTERPRO: IPR002012;
CC PFAM: PF00446; GNRH: 1.
DR PROSITE: PS00473; GNRH: 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
FT NON_TER 1 1
FT CHAIN 1 >63 PROGNADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >63 GNRH-ASSOCIATED PEPTIDE I (BY
FT ACT_SITE 3 3 SIMILARITY).
FT ACTIVITY (BY SIMILARITY).
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (BY

```

FT MOD_RES 10 10 SIMILARITY.
 FT AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY
 FT NON_PER 63 63 SIMILARITY).
 FT SEQUENCE 63 AA: 7370 MM: FC94995676F77180 CRC64;

Query Match 55.6%; Score 60; DB 1; Length 63;
 Best Local Similarity 90.0%; Pred. No. 0.058;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 8 EHWSTGLRPG 17
 :|||||
 Db 1 QHWSYGLRPG 10

RESULT 10
 GONL_CAVPO STANDARD; PRT: 92 AA.

AC 054713;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
 DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
 DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
 GN GNRH1 OR GNRH OR LHRH.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HARTLEY WHITE; TISSUE-HYPOTHALAMUS;
 RX MEDLINE-97462693; PubMed-9322920;
 RA Jimenez-Linan M., Rubin B.S., King J.C.;
 RT "Examination of guinea pig luteinizing hormone-releasing hormone gene
 RT reveals a unique decapeptide and existence of two transcripts in the
 RT brain.";
 RL Endocrinology 138:4123-4130(1997).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
 CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
 CC HORMONES.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF033346; AAB87688.1; -
 CC DR INTERPRO: IPR002012; -
 CC DR PROSITE: PS00473; GNRH.1.
 CC KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 CC Placenta; Signal.

FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 92 PROGNADOLIBERIN I.
 FT PEPTIDE 24 33 GNRH-ASSOCIATED PEPTIDE I.
 FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
 FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
 FT ACTIVITY (BY SIMILARITY).
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
 FT MOD_RES 33 33 SIMILARITY).
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
 FT SEQUENCE 92 AA: 10279 MM: ACF74613F456D663 CRC64;

Query Match 52.8%; Score 57; DB 1; Length 92;
 Best Local Similarity 52.9%; Pred. No. 0.21;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 CPPPSEHWSYGLRPG 17
 :|||||
 Db 17 CVENGSGYWSYGLRPG 33

RESULT 11
 GONL_CHICK STANDARD; PRT: 92 AA.

AC P37042; P20407;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I),
 DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
 DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WHITE LECHORN;
 RX MEDLINE-94059355; PubMed-7902095;
 RA Dunn I.C., Chen Y., Hook C., Sharp P.J., Sang H.M.;
 RT "Characterization of the chicken preprogonadotrophin-releasing
 RT hormone-I gene.";
 RL J. Mol. Endocrinol. 11:19-29(1993).
 RN [2]
 RP SEQUENCE OF 24-33.
 RC TISSUE-HYPOTHALAMUS;
 RX MEDLINE-82265778; PubMed-7050119;
 RA King J.A., Millar R.P.;
 RT "Structure of chicken hypothalamic luteinizing hormone-releasing
 RT hormone. II. Isolation and characterization.";
 RL J. Biol. Chem. 257:10729-10732(1982).
 RN [3]
 RP SEQUENCE OF 24-33.
 RC TISSUE-HYPOTHALAMUS;
 RA King J.A., Millar R.P.;
 RT "Structure of avian hypothalamic gonadotrophin-releasing hormone.";
 RL S. Afr. J. Sci. 78:124-125(1982).
 RN [4]
 RP SYNTHESIS OF 24-33.
 RX MEDLINE-82265777; PubMed-7050118;
 RA King J.A., Millar R.P.;
 RT "Structure of chicken hypothalamic luteinizing hormone-releasing
 RT hormone. I. Structural determination on partially purified
 RT material.";
 RL J. Biol. Chem. 257:10722-10728(1982).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X69491; CAA49246.1; -
 CC DR PIR: S33507; S33507.
 CC DR INTERPRO: IPR002012; -
 CC DR PFAM: PF00446; GNRH.1.
 CC DR PROSITE: PS00473; GNRH.1.
 CC KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 CC Signal.

FT SIGNAL 1 23 PROGNADOLIBERIN I.
 FT CHAIN 24 92 GNRH-ASSOCIATED PEPTIDE I.
 FT PEPTIDE 24 33 GNRH-ASSOCIATED PEPTIDE I.
 FT PEPTIDE 37 92 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 24 24

FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE 92 AA: 10206 MW: 61AE87EBAF508BEA CRC64;

Query Match 52.8%; Score 57; DB 1; Length 92;
Best Local Similarity 72.7%; Pred. No. 0.21;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17
DB 23 AOHWSYGLQPG 33

RESULT 12

CON1_ALIMI STANDARD; PRT; 10 AA.
AC P37041; P20407;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GONADOLIBERIN I (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I) (LH-RH I)
DE (LULIBERIN I).
OS Alligator mississippiensis (American alligator).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Crocodylidae; Alligatorinae; Alligator.
RN [1]
RP SEQUENCE.
RC TISSUE-BRAIN;
RX MEDLINE-9152338; PubMed-1882082;
RA Lofjoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
*Primary structure of the two forms of gonadotropin-releasing hormone from brains of the American alligator (Alligator mississippiensis).
RL Regul. Pept. 33:105-116(1991).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC PIR: A60066; RHAQI.
DR INTERPRO: IPR002012; -
DR PFAM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Hormone; Amidation; Hypothalamus.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA: 1172 MW: 284B23D7286B45A3 CRC64;

Query Match 51.9%; Score 56; DB 1; Length 10;
Best Local Similarity 80.0%; Pred. No. 0.031;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 8 EHSYGLRPG 17
DB 1 OHWSYGLQPG 10

RESULT 13

CON1_HAPBU STANDARD; PRT; 94 AA.
ID GON1_HAPBU
AC P51918; 093387;
DT 01-OCT-1996 (Rel. 34, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I) (LH-RH I) (LULIBERIN I).
DE (LH-RH I) (LULIBERIN I).
GN GNRH1.
OS Haplochromis burtoni.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;
OC Cichlidae; Astacotilapia.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95396797; PubMed-7667296;

RA White S.A., Kasten T.L., Bond C.T., Adelman J.P., Fernald R.D.;
RT "Three gonadotropin-releasing hormone genes in one organism suggest
RT novel roles for an ancient peptide."
RL Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99061842; PubMed-9843638;
RA White R.B., Fernald R.D.;
RT "Ontogeny of gonadotropin-releasing hormone (GNRH) gene expression
RT reveals a distinct origin for GNRH-containing neurons in the
RT midbrain."
RL Gen. Comp. Endocrinol. 112:322-329(1998).
RN [3]
RP SEQUENCE OF 23-32.
RC TISSUE-PITUITARY;
RX MEDLINE-95372591; PubMed-7644702;
RA Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,
RA White S.A., Francis R.C., Fernald R.D., Licht P., Wardy C.,
RA Sherwood N.M.;
RT "Primary structure of solitary form of gonadotropin-releasing hormone (GNRH) in cichlid pituitary; three forms of GNRH in brain of cichlid
RT and pumpkinseed fish."
RL Regul. Pept. 57:43-53(1995).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE
CC RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALMIC-PITUITARY-
CC GONADAL AXIS.
CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS
CC TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPOTHALMIC AXONS.
CC -1- MASS SPECTROMETRY: MW-1113.9; METHOD-MALDI; RANGE-23-32.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U31865; AAC59691.1; -
DR EMBL: AF076961; AAC27716.1; -
DR INTERPRO: IPR002012; -
DR PFAM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Multigene family.
FT CHAIN 1 22 PROGONADOLIBERIN I.
FT PEPTIDE 23 32 GONADOLIBERIN I.
FT PEPTIDE 36 94 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
FT MOD_RES 23 23 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 32 32 AMIDATION (G-33 PROVIDE AMIDE GROUP).
FT CONFLICT 86 94 ENGHRTFKK -> KMDTGSRNRERFL (IN REF. 1).
SQ SEQUENCE 94 AA: 10382 MW: E57DBA8333278D7 CRC64;

Query Match 51.9%; Score 56; DB 1; Length 94;
Best Local Similarity 60.0%; Pred. No. 0.3;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 PPSSEHWSYGLRPG 17
DB 18 PQGCCOHWSYGLSPG 32

RESULT 14

CON1_CLAGA STANDARD; PRT; 80 AA.
ID GON1_CLAGA
AC P34339;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)

```

DE (LH-RH) (LULIBERIN I).
OS Clarias gariepinus (Sharptooth catfish) (African catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Siluriformes; Clariidae; Clarias.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE-94291651; PubMed-8020492;
RA Bogerd J., Zandbergen T., Andersson E., Goos H.;
RT "Isolation, characterization and expression of cDNAs encoding the
RT catfish-type and chicken-II-type gonadotropin-releasing-hormone
RT precursors in the African catfish.";
RL Eur. J. Biochem. 222:541-549(1994).
RN [2]
RP SEQUENCE OF 25-31.
RC TISSUE-BRAIN;
RX MEDLINE-92392313; PubMed-1520292;
RA Bogerd J., Li K.W., Janssen-Dommerholt C., Goos H.;
RT "Two gonadotropin-releasing hormones from African catfish (Clarias
RT gariepinus)";
RL Biochem. Biophys. Res. Commun. 187:127-134(1992).
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X78048; CAA54970.1; -.
DR EMBL: X78049; CAA54971.1; -.
DR PIR: JC1242; RHIDIS.
DR INTERPRO: IPR002012; -.
DR PFAM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 21
FT CHAIN 22 80 PROGONADOLIBERIN I.
FT PEPTIDE 22 31 GONADOLIBERIN I.
FT PEPTIDE 35 80 GNRH-ASSOCIATED PEPTIDE I.
FT MOD_RES 22 22 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 31 31 AMIDATION (G-32 PROVIDE AMIDE GROUP).
FT VARIANT 47 47 S -> R.
FT VARIANT 60 60 G -> R.
SQ SEQUENCE 80 AA; 8893 MW; 0BE5EE0F4FF861A CRC64;

```

Query Match 50.9%; Score 55; DB 1; Length 80;

Best Local Similarity 66.7%; Pred. No. 0.34; Mismatches 1; Indels 0; Gaps 0;

QY 6 SEHWSYGLRPG 17
 1:|||||11
 DB 20 SAQWHSGLNPG 31

```

RESULT 15
GONI PAGMA
ID GONI PAGMA STANDARD; PRT: 95 AA.
AC P70074;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE (LH-RH I) (LULIBERIN I).
OS Pagrus major (Red sea bream) (Chrysophrys major).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

```

```

OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Chrysophrys.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Okuzawa K., Graneman J., Bogerd J., Goos H., Zohar Y., Kagawa H.;
RT Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
CC SIMILARITY).
CC -----
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D86582; BAA13129.1; -.
DR INTERPRO: IPR002012; -.
DR PFAM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Multigene family.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 95 PROGONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 95 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP)
FT (BY SIMILARITY).
SQ SEQUENCE 95 AA; 10566 MW; 61E79C990328D73E CRC64;

```

Query Match 50.0%; Score 54; DB 1; Length 95;

Best Local Similarity 80.0%; Pred. No. 0.56; Mismatches 1; Indels 0; Gaps 0;

QY 8 EHWWSYGLRPG 17
 1:|||||11
 DB 24 QWWSYGLSPG 33

Search completed: March 2, 2001, 11:01:19
 Job time: 434 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:55:38 ; Search time 56.87 Seconds
(without alignments)
20.297 Million cell updates/sec

Title: US-09-306-689-3

Perfect score: 108

Sequence: 1 CPPPSEHWSYGLRPG 17

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 6790655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	63.9	92	1	RHHUG
2	68	63.0	67	2	178541
3	68	63.0	90	1	RHMSG
4	68	63.0	92	1	RHRTG
5	61	56.5	89	2	151423
6	60	55.6	10	1	RHPGG
7	60	55.6	10	1	RHSHG
8	57	52.8	92	2	150644
9	56	51.9	10	1	RHAQ1
10	55	50.9	98	1	150739
11	55	50.9	80	1	RHRTD
12	52	48.1	2357	2	AS9249
13	51	47.2	74	2	151092
14	51	47.2	82	2	151180
15	51	47.2	82	2	151355
16	51	47.2	82	2	151365
17	51	47.2	82	2	151331
18	51	47.2	90	2	A23735
19	51	47.2	90	2	151095
20	50	46.3	497	2	115812
21	49.5	44.8	437	2	C48274
22	47.5	44.0	305	2	T02851
23	47	43.5	10	2	A21114
24	47	43.5	190	2	T29068
25	47	43.5	226	2	T34808
26	47	43.5	322	2	S25299
27	47	43.5	325	2	T01830
28	47	43.5	327	2	T33884
29	47	43.5	361	2	S19552

30	47	43.5	396	2	A36339	FL0 protein - gard
31	47	43.5	769	2	A42073	potassium channel
32	47	43.5	769	2	156546	Shaw type potassium
33	46	42.6	86	1	RHID25	gonadoliblerin II p
34	46	42.6	164	2	C82629	hypothetical prote
35	46	42.6	252	2	T45737	hypothetical prote
36	46	42.6	404	2	T32725	hypothetical prote
37	46	42.6	412	2	B44418	surface antigen -
38	46	42.6	543	2	H70891	probable acid--COA
39	46	42.6	551	2	A60047	adenylyl cyclase-a
40	46	42.6	756	1	B28821	1-phosphatidylinos
41	45	41.7	122	2	A05175	hypothetical prote
42	45	41.7	164	2	G71427	hypothetical prote
43	45	41.7	169	2	T34520	hypothetical prote
44	45	41.7	282	2	C72282	oligopeptide ABC t
45	45	41.7	335	2	D83142	hypothetical prote

ALIGNMENTS

RESULT 1

RHHUG
gonadoliblerin precursor - human
N:Alternate names: gonadotropin releasing hormone (GNRH); luteinizing hormone releas
N:Contains: gonadoliblerin-associated protein (GAR); progadoliblerin
C:Species: Homo sapiens (man)
C:Date: 17-Mar-1987 #sequence-revision 21-Jul-1995 #text-change 18-Jun-1999
C:Accession: S05308; A26173; A93342; A90108; A01410; S45718
R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
Nucleic Acids Res. 17, 6403-6404, 1989
A:Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone
A:Reference number: S05308; MUID: 8936682
A:Accession: S05308
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-92 <NAV>
A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gona
A:Reference number: A94090; MUID: 86094338
A:Accession: A26173
A:Molecule type: mRNA
A:Residues: 1-92 <ADE>
A:Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749
A:Experimental source: hypothalamus
R:Seeburg, P.H.; Adelman, J.P.
Nature 311, 666-668, 1984
A:Title: Characterization of cDNA for precursor of human luteinizing hormone releasin
A:Reference number: A93342; MUID: 85012739
A:Accession: A93342
A:Molecule type: mRNA
A:Residues: 1-15, 'S', 17-92 <SEE>
A:Cross-references: GB:X01059; NID:g34356; PIDN:CAA5526.1; PID:g34357
A:Experimental source: placenta
R:Tan, L.; Kousseau, P.
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982
A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized
A:Reference number: A90108; MUID: 83126573
A:Accession: A90108
A:Molecule type: protein
A:Residues: 24-33 <TAN>
A:Experimental source: placental trophoblasts
R:Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amst
FEBS Lett. 346, 203-206, 1994
A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by
A:Reference number: S45718; MUID: 9428397
A:Contents: annotation; degradation pathway of synthetic hormone
C:Genetics:
A:Gene: GDB:GNRH; LHRH; GRH
A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760
A:Map position: 8p21-8p11.2

A; Introns: 47/3; 79/3
C; Function:
A; Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitropin
A; Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C; Superfamily: gonadoliberin
F; 1-23/Domain: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglyutamic acid
F; 1-23/Domain: signal sequence #status predicted <SIG>
F; 24-92/Product: progadoliberin #status predicted <PGN>
F; 24-92/Product: gonadoliberin #status experimental <MAT>
F; 37-92/Product: gonadoliberin-associated protein #status predicted <GAP>
F; 24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F; 33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 63.9%; Score 69; DB 1; Length 92;
Best Local Similarity 70.6%; Pred. No. 0.01;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CPPPSSEHWSYGLRPG 17
11:|||||
Db 17 CVEGSSQHWSYGLRPG 33

RESULT 2
178541
gonadoliberin precursor - rhesus macaque (fragment)
N; Alternate names: luteinizing hormone releasing hormone
C; Species: Macaca mulatta (rhesus macaque)
C; Date: 02-Aug-1996 #sequence; revision 02-Aug-1996 #text-change 16-Jul-1999
C; Accession: 178541
R; Ma, Y.J.; Costa, M.E.; Ojeda, S.R.
Neuroendocrinology 60, 346-359, 1994
A; Title: Developmental expression of the genes encoding transforming growth factor alpha
A; Reference number: I58134; MUID:95124501
A; Accession: 178541
A; Status: preliminary; translated from GB/EMBL/DBJ
A; Molecule type: mRNA
A; Residues: 1-67 <RES>
A; Cross-references: GB:S75918; NID:g912831; PIDN:AAB33096.1; PID:g912832
C; Superfamily: gonadoliberin

Query Match 63.0%; Score 68; DB 2; Length 67;
Best Local Similarity 91.7%; Pred. No. 0.01;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 SSEHWSYGLRPG 17
11:|||||
Db 4 SSOHWSYGLRPG 15

RESULT 3
RHMSG
gonadoliberin precursor - mouse
N; Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing
N; Contains: gonadoliberin; gonadoliberin-associated protein (GAP)
C; Species: Mus musculus (house mouse)
C; Date: 31-Dec-1993 #sequence; revision 18-Mar-1997 #text-change 18-Jun-1999
C; Accession: A47578
R; Mason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikolics,
Science 234, 1366-1371, 1986
A; Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible fo
A; Reference number: A47578; MUID:87069928
A; Accession: A47578
A; Molecule type: DNA
A; Residues: 1-90 <MAS>
A; Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387175
C; Genetics:
A; Introns: 45/3; 77/3
C; Function:
A; Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitropin
A; Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C; Superfamily: gonadoliberin

C; Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglyutamic acid
F; 1-23/Domain: signal sequence #status predicted <SIG>
F; 23-31/Product: gonadoliberin #status predicted <GAB>
F; 35-90/Product: gonadoliberin-associated protein #status predicted <GAP>
F; 22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F; 31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 63.0%; Score 68; DB 1; Length 90;
Best Local Similarity 91.7%; Pred. No. 0.014;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 SSEHWSYGLRPG 17
11:|||||
Db 20 SSOHWSYGLRPG 31

RESULT 4
RHRTG
gonadoliberin precursor - rat
N; Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hor
N; Contains: gonadoliberin; prolactin release-inhibiting factor
C; Species: Rattus norvegicus (Norway rat)
C; Date: 31-Mar-1988 #sequence; revision 31-Mar-1988 #text-change 18-Jun-1999
C; Accession: A40147; B26173; A48410
R; Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
Mol. Endocrinol. 3, 1257-1262, 1989
A; Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic
A; Reference number: A40147; MUID:89384661
A; Accession: A40147
A; Molecule type: DNA
A; Residues: 1-92 <BON>
A; Cross-references: GB:M31670; NID:g204447; PIDN:AAA1264.1; PID:g204448
R; Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A; Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gona
A; Reference number: A94090; MUID:86094338
A; Accession: B26173
A; Molecule type: mRNA
A; Residues: 1-92 <ADE>
A; Cross-references: GB:M12579; NID:g204445; PIDN:AAA1263.1; PID:g204446
R; Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blalock, J.E.
Cell. Mol. Neurobiol. 12, 447-454, 1992
A; Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing horm
A; Reference number: A48410; MUID:93105480
A; Accession: A48410
A; Status: preliminary
A; Molecule type: mRNA
A; Residues: 1-92 <MAI>
A; Cross-references: GB:S50870; NID:g262059; PIDN:AAB24572.1; PID:g262060
A; Experimental source: thymus
A; Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBI:P:121083)
C; Genetics:
A; Introns: 47/3; 79/3
C; Function:
A; Description: stimulates pituitary secretion of luteotropin and follitropin
A; Note: gonadoliberin-associated protein may have prolactin release inhibiting activi
C; Superfamily: gonadoliberin
C; Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglyutamic acid
F; 1-23/Domain: signal sequence #status predicted <SIG>
F; 24-92/Product: progadoliberin #status predicted <PGN>
F; 24-92/Product: gonadoliberin #status predicted <GAB>
F; 37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>
F; 24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F; 33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 63.0%; Score 68; DB 1; Length 92;
Best Local Similarity 91.7%; Pred. No. 0.014;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 SSEHWSYGLRPG 17
11:|||||

Db 22 SSOHWSYGLRPG 33

RESULT 5
151423
gonadoliberin precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: 151423
R:Hayes, W.P.; Wray, S.; Battey, J.F.
Endocrinology 134, 1835-1845, 1994
A:Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved domain
A:Reference number: 151423; MUID:94185563
A:Accession: 151423
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-89 <RAY>
A:Cross-references: GB:L28040; NID:g496291; PIDN:AAA9728.1; PID:g496292
C:Genetics:
A:Gene: GnRH-I
C:Superfamily: gonadoliberin

Query Match 56.5%; Score 61; DB 2; Length 89;
Best Local Similarity 81.8%; Pred. No. 0.12;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 7 SEHWSYGLRPG 17
:|||||
Db 23 AQHWSYGLRPG 33

RESULT 6
RHPG
gonadoliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 13-Jul-1991 #sequence_revision 13-Jul-1991 #text_change 18-Mar-1997
C:Accession: A01411
R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of the
A:Reference number: A90172; MUID:72114303
A:Accession: A01411
A:Molecule type: protein
A:Residues: 1-10 <BAB>
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase method
A:Reference number: A90176; MUID:72065376
A:Contents: annotation
A:Note: the synthetic and natural hormones have the same physicochemical and biological
R:Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A:Reference number: A90175; MUID:72117544
A:Contents: annotation
A:Note: Trp-3 appears to be essential for biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle-stimulating hormone
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 55.6%; Score 60; DB 1; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.016;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 8 EHWYGLRPG 17
:|||||
Db 1 QHWSYGLRPG 10

RESULT 7

RHSHG
gonadoliberin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C:Accession: A93780; A01411
R:Burgus, R.; Butcher, M.; Amoss, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing factor
A:Reference number: A93780; MUID:72094314
A:Accession: A93780
A:Molecule type: protein
A:Residues: 1-10 <BUR>
A:Note: The natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle-stimulating hormone
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 55.6%; Score 60; DB 1; Length 10;
Best Local Similarity 90.0%; Pred. No. 0.016;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 8 EHWYGLRPG 17
:|||||
Db 1 QHWSYGLRPG 10

RESULT 8
150644
gonadoliberin I precursor - chicken
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Gallus gallus (chicken)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
C:Accession: I50644; S33507
R:Dun, I.C.; Chen, Y.; Hook, C.; Sharp, P.J.; Sang, H.M.
J. Mol. Endocrinol. 11, 19-29, 1993
A:Title: Characterization of the chicken preprogonadotropin-releasing hormone-I gene
A:Reference number: I50644; MUID:94059355
A:Accession: I50644
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-92 <DU2>
A:Cross-references: EMBL:X69491; NID:g496326; PIDN:CAA49246.1; PID:g311612
C:Genetics:
A:Insertions: 47/3; 79/3
C:Superfamily: gonadoliberin

Query Match 52.8%; Score 57; DB 2; Length 92;
Best Local Similarity 72.7%; Pred. No. 0.41;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
OY 7 SEHWSYGLRPG 17
:|||||
Db 23 AQHWSYGLRPG 33

RESULT 9
RHA01
gonadoliberin I - American alligator
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: A60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swan
Regul. Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains
A:Reference number: A60066; MUID:91352338
A:Accession: A60066

A:Molecule type: protein
 A:Residues: 1-10 <LOV>
 C:Superfamily: gonadoliberin
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 51.9%; Score 56; DB 1; Length 10;
 Best Local Similarity 80.0%; Pred. No. 0.054;
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 8 EHMSYGLRPG 17
 :|||||:|
 Db 1 OHMSYGLPG 10

RESULT 10
 150739

gonadotropin-releasing hormone - Cichlid (Haplochromis burtoni)
 C:Species: Haplochromis burtoni

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: 150739

R:White, S.A.; Kasten, T.L.; Bond, C.T.; Adelman, J.P.; Fernald, R.D.
 Proc. Natl. Acad. Sci. U.S.A. 92, 8363-8367, 1995

A:Title: Three gonadotropin-releasing hormone genes in one organism suggest novel roles
 A:Reference number: 150739; MUID:95396797

A:Accession: 150739

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-98 <WH>

A:Cross-references: EMBL:U01865; NID:9905398; PIDN:MAC59691.1; PID:9905399

C:Superfamily: gonadoliberin

Query Match 51.9%; Score 56; DB 2; Length 98;
 Best Local Similarity 60.0%; Pred. No. 0.6;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 PPSEHMSYGLRPG 17
 :|||||:|
 Db 18 POGCCMHSTGLSPG 32

RESULT 11
 RHID15

gonadoliberin I precursor - sharpshooth catfish

N:Alternate names: gonadoliberin, catfish-type; gonadotropin-releasing hormone I (GnRH-I)

N:Contains: gonadoliberin I; gonadoliberin I-associated protein form I; gonadoliberin I-

C:Species: *Clarias gariepinus* (sharpshooth catfish)

C>Date: 30-Sep-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999

C:Accession: S45602; J01242; S42936; S42937

R:Boigerd, J.; Zandbergen, T.; Andersson, E.; Goos, H.
 Eur. J. Biochem. 222, 541-549, 1994

A:Title: Isolation, characterization and expression of cDNAs encoding the catfish-type A

A:Reference number: S45600; MUID:94291651

A:Molecule type: mRNA

A:Residues: 1-80 <BOG1>

A:Cross-references: EMBL:X78049; NID:9459433; PIDN:CA54971.1; PID:9459434

A:Note: gonadoliberin I-associated protein form I

A:Molecule type: mRNA

A:Accession: S45601

A:Residues: 1-46, 'S', 48-59, 'G', 61-80 <BOG2>

A:Cross-references: EMBL:X78049; NID:9459433; PIDN:CA54970.1; PID:9459432

A:Note: gonadoliberin I-associated protein form II, presumed to be a polymorphic form

R:Boigerd, J.; Li, K.W.; Janssen-Dommerholt, C.; Goos, H.
 Biochem. Biophys. Res. Commun. 187, 127-134, 1992

A:Title: Two gonadotropin-releasing hormones from African catfish (*Clarias gariepinus*).
 A:Reference number: J01242; MUID:92392313

A:Accession: J01242

A:Molecule type: protein

A:Residues: 22-31 <BOG3>

A:Experimental source: brain
 C:Superfamily: gonadoliberin
 C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-31/Product: gonadoliberin I #status experimental <MAT1>
 F:33-80/Product: gonadoliberin I-associated protein #status predicted <MAT2>
 F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi
 F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 50.9%; Score 55; DB 1; Length 80;
 Best Local Similarity 66.7%; Pred. No. 0.66;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 6 SSEHMSYGLRPG 17
 :|||||:|
 Db 20 SAQHMSHGLNPG 31

RESULT 12
 A59249

Class VII unconventional myosin - slime mold (*Dictyostelium discoideum*)

C:Species: *Dictyostelium discoideum*

C>Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 08-Sep-2000

C:Accession: A59249

R:Titus, M.A.; Kuspa, A.; Loomis, W.F.
 Proc. Natl. Acad. Sci. U.S.A. 91, 9446-9450, 1994

A:Title: Discovery of myosin genes by physical mapping in *Dictyostelium*.
 A:Reference number: Z20873; MUID:95023928

A:Accession: A59249

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-2357 <TT>

A:Cross-references: GB:U35321; NID:96226760; PIDN:AAF06035.1; PID:96226761

A:Experimental source: gene myoI; product class VII unconventional myosin

R:Titus, M.A.
 Curr. Biol. 9, 1297-1303, 1999

A:Title: A class VII unconventional myosin is required for phagocytosis.

A:Reference number: A59249

A:Contents: annotation

C:Genetics:

A:Gene: myoI

C:Superfamily: myosin motor domain homology

F:16-676/Domain: myosin motor domain homology <MMO>

Query Match 48.1%; Score 52; DB 2; Length 2357;
 Best Local Similarity 60.0%; Pred. No. 59;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 PPPSEHMSYGLRPG 16
 :|||||:|
 Db 1027 PPPSSSSMDGLRPG 1041

RESULT 13
 I51092

gonadotropin releasing hormone - chinook salmon (fragment)

C:Species: *Oncorhynchus tshawytscha* (chinook salmon)

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Aug-1997

C:Accession: I51092

R:Klungland, H.; Lorens, J.B.; Andersen, O.; Kisen, G.O.; Alestrom, P.
 Mol. Cell. Endocrinol. 84, 167-174, 1992

A:Title: The Atlantic salmon prepro-gonadotropin releasing hormone gene and mRNA.

A:Reference number: I51040; MUID:92267241

A:Accession: I51092

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-74 <KL>

A:Cross-references: EMBL:X79711; NID:9499322; PID:9499323

C:Genetics:

A:Gene: GnRH

A:Introns: 38/3; 65/3

Job time: 259 sec

Query Match 47.2%; Score 51; DB 2; Length 74;
 Best Local Similarity 72.7%; Pred. No. 2.1;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17
 1:|||||
 DB 15 SQHWSYGLRPG 25

RESULT 14

151180

gonadotropin-releasing hormone - cherry salmon

C:Species: Oncorhynchus masou (cherry salmon)

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Aug-1997

C:Accession: 151180

R:Suzuki, M.; Hyodo, S.; Kobayashi, M.; Aida, K.; Urano, A.

J. Mol. Endocrinol. 9, 73-82, 1992.

A:Title: Characterization and localization of mRNA encoding the salmon-type gonadotrophin

A:Reference number: 151180; MUID:92284893

A:Accession: 151180

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-82 <SU2>

A:Cross-references: GB:S44614; NID:9254824; PID:9254825

Query Match 47.2%; Score 51; DB 2; Length 82;
 Best Local Similarity 72.7%; Pred. No. 2.3;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17
 1:|||||
 DB 23 SQHWSYGLRPG 33

RESULT 15

151355

gonadotropin releasing hormone - Atlantic salmon

C:Species: Salmo salar (Atlantic salmon)

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: 151355

R:Klungland, H.; Lorens, J.B.; Andersen, O.; Kisen, G.O.; Alestrom, P.

Mol. Cell. Endocrinol. 84, 167-174, 1992.

A:Title: The Atlantic salmon prepro-gonadotropin releasing hormone gene and mRNA.

A:Reference number: 151040; MUID:92267241

A:Accession: 151355

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-82 <KLU>

A:Cross-references: EMBL:X79709; NID:9499341; PID:9499342

A:Accession: 151355

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-82 <KLU>

A:Cross-references: EMBL:X74957; NID:9402786; PIDN:CA52912.1; PID:9402787

C:Genetics:

A:Gene: GnRH

A:introns: 46/3; 73/3

Query Match 47.2%; Score 51; DB 2; Length 82;
 Best Local Similarity 72.7%; Pred. No. 2.3;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 SEHWSYGLRPG 17
 1:|||||
 DB 23 SQHWSYGLRPG 33

Search completed: March 2, 2001, 10:55:39

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:53:10 ; Search time 97.13 Seconds
(without alignments)
16.894 Million cell updates/sec

Title: US-09-306-689-2
Perfect score: 58
Sequence: 1 XHMSYXLRGEXXX 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_15:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	87.9	91	13	Q9PRH0
2	45	77.6	87	13	Q9Y126
3	45	77.6	95	13	Q73812
4	45	77.6	99	13	Q91A10
5	43	74.1	90	13	Q91A02
6	42	72.4	565	2	Q9KM11
7	39	67.2	508	10	Q82588
8	39	67.2	828	10	Q9SZR5
9	38	65.5	33	13	Q9W7G0
10	38	65.5	33	13	Q9PT34
11	38	65.5	82	13	Q92094
12	38	65.5	82	13	Q9W7G1
13	38	65.5	82	13	Q91B00
14	38	65.5	82	13	Q91B09
15	38	65.5	88	13	Q9PSY9
16	38	65.5	90	13	Q91A09
17	38	65.5	205	2	Q9S299
18	37	63.8	954	10	Q22613
19	37	63.8	965	10	Q9M4N4

20	37	63.8	966	10	Q9M4N3	Q9M4N3 medicago tr
21	36	62.1	289	2	Q9X2P4	Q9X2P4 mycobacteri
22	36	62.1	321	10	Q9S0G4	Q9S0G4 papaver som
23	36	62.1	379	10	Q9LV12	Q9LV12 arabidopsis
24	36	62.1	425	5	Q9U9U4	Q9U9U4 drosophila
25	36	62.1	612	10	Q9M8N3	Q9M8N3 arabidopsis
26	36	62.1	954	10	Q9SMW2	Q9SMW2 nicotiana p
27	36	62.1	954	10	Q9SAJ1	Q9SAJ1 arabidopsis
28	36	62.1	956	10	Q9LV11	Q9LV11 arabidopsis
29	35	60.3	32	5	Q9U340	Q9U340 caenorhabdi
30	35	60.3	82	1	Q28696	Q28696 archaeoglob
31	35	60.3	148	2	Q51920	Q51920 streptomyce
32	35	60.3	155	3	Q9UR56	Q9UR56 kluyveromyc
33	35	60.3	155	8	Q9XMD4	Q9XMD4 kluyveromyc
34	35	60.3	161	2	Q9RVDB	Q9RVDB deinococcus
35	35	60.3	220	5	Q21256	Q21256 caenorhabdi
36	35	60.3	479	2	Q9LON5	Q9LON5 streptomyce
37	35	60.3	485	2	Q9LON3	Q9LON3 streptomyce
38	35	60.3	526	2	Q9J0A1	Q9J0A1 chlamydia p
39	35	60.3	567	4	Q13208	Q13208 homo sapien
40	35	60.3	691	10	Q23482	Q23482 arabidopsis
41	35	60.3	711	4	Q14870	Q14870 homo sapien
42	35	60.3	711	4	Q14870	Q14870 homo sapien
43	35	60.3	1072	6	Q77690	Q77690 bos taurus
44	35	60.3	1232	10	Q9M264	Q9M264 arabidopsis
45	34	58.6	119	1	Q9Y9R2	Q9Y9R2 aetopyrum p

ALIGNMENTS

RESULT 1	Q9PRH0	PRELIMINARY;	PRT;	91 AA.
AC	Q9PRH0;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DI	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DI	01-JUN-2000 (TREMBLrel. 14, Last annotation update)			
DE	PREPRO-MGNRH PRECURSOR.			
OS	Anguilla japonica (Japanese eel).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillioidei;			
OX	NCBI_TaxID=7937;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BRAIN;			
RA	Okubo K., Suetake H., Aida K.;			
RT	"Expression of two gonadotropin-releasing hormone (GnRH) precursor genes in various tissues of the Japanese eel and evolution of GnRH.;"			
RL	Zool. Sci. 16:471-478(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Okubo K., Suetake H., Aida K.;			
RT	"A splicing variant for the prepro-mammalian gonadotropin-releasing hormone (prepro-mgnrh) mRNA is present in the brain and various peripheral tissues of the Japanese eel.;"			
RL	Zool. Sci. 16:645-651(1999).			
DR	EMBL: AB026989; BAA82608.1; -			
DR	EMBL: AB026991; BAA83597.1; -			
DR	INTERPRO: IPR002012; -			
DR	PFAM: PF00446; GNRH.1.			
DR	PROSITE: PS00473; GNRH.1.			
KW	SIGNAL.			
FT	CHAIN 1 22	POTENTIAL.		
FT	CHAIN 23 32	MGNRH.		
FT	CHAIN 33 91	GNRH ASSOCIATED PEPTIDE.		
FT	SEQUENCE 91 AA; 9893 MW; BAI5C9DC0843A7B CRC64;			
Query Match	87.9%;	Score 51;	DB 13;	Length 91;
Best Local Similarity	88.9%;	Pred. No. 0.023;		
Matches	8;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
 |||||
 Db 24 HWSYGLSPG 32

RESULT 2
 ID 09Y126 PRELIMINARY; PRT; 87 AA.
 AC 09Y126;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN) (FRAGMENT).
 OS Sparus aurata (Gillthead sea bream).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae; Sparidae; Sparus.
 OX NCBL_TaxID=8175;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=OVARY;
 RA Nabissi M.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND POLICIE- STIMULATING HORMONES.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 DR EMBL; AF046801; AAD02427.1; -;
 DR INTERPRO; IPR002012; -;
 DR PFAM; PF00446; GNRH; 1.
 DR PROSITE; PS00473; GNRH; 1.
 KM Hormone; Amidation.
 FT NON_TER 1 87
 FT NON_TER 87 87
 SQ SEQUENCE 87 AA; 9871 MW; 0D246353D96782A CRC64;

Query Match 77.6%; Score 45; DB 13; Length 87;
 Best Local Similarity 77.8%; Pred. No. 0.28;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
 |||||
 Db 22 HWSYGLSPG 30

RESULT 3
 ID 073812 PRELIMINARY; PRT; 95 AA.
 AC 073812;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN).
 OS More saxatilis (Striped bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae; Moronidae; Morone.
 OX NCBL_TaxID=34816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Chow M.M., Kight K.E., Gotthif Y., Alok D., Zohar Y.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND POLICIE- STIMULATING HORMONES.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 DR EMBL; AF056314; AAD03817.1; -;
 DR INTERPRO; IPR002012; -;
 DR PFAM; PF00446; GNRH; 1.
 DR PROSITE; PS00473; GNRH; 1.

DR PRODOM; PD005581; -; 1.
 KM Hormone; Amidation.
 SQ SEQUENCE 95 AA; 10411 MW; 980C698BFC279BFC CRC64;

Query Match 77.6%; Score 45; DB 13; Length 95;
 Best Local Similarity 77.8%; Pred. No. 0.31;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
 |||||
 Db 24 HWSYGLSPG 32

RESULT 4
 ID 091A10 PRELIMINARY; PRT; 99 AA.
 AC 091A10;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE GONADOTROPIN-RELEASING HORMONE SEABREAM ISOFORM.
 OS Dicentrarchus labrax (European sea bass).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae; Moronidae; Dicentrarchus.
 OX NCBL_TaxID=13489;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RA Gonzalez-Martinez D., Madigou T., Zmora N., Anglade I., Zanny S., Zohar Y., Elizur A., Munoz-Cueto J.A., Kah O.;
 RL "Differential expression of three different prepro-GNRH (gonadotrophin-releasing hormone) messengers in the brain of the European sea bass (Dicentrarchus labrax)."
 RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RA Zmora N., Zohar Y., Elizur A.;
 RT "3 GNRH form in the seabass Dicentrarchus labrax."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF224279; AAF62898.1; -;
 SQ SEQUENCE 99 AA; 10758 MW; EC8AEEC93CC02904 CRC64;

Query Match 77.6%; Score 45; DB 13; Length 99;
 Best Local Similarity 77.8%; Pred. No. 0.32;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
 |||||
 Db 28 HWSYGLSPG 36

RESULT 5
 ID 091A02 PRELIMINARY; PRT; 90 AA.
 AC 091A02;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE GONADOTROPIN-RELEASING HORMONE.
 OS Rana dybowskii (Frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
 OX NCBL_TaxID=71582;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RA Yoo M.S., Kang H.M., Choi H.S., Chun S.Y., Troskie B., Millar R.P., Kwon H.B.;

RT "Molecular Cloning, Distribution and Pharmacological Characterization
of a Novel Gonadotropin-Releasing Hormone([trp8]GnRH) in Frog Brain.";
RL Mol. Cell. Endocrinol. 0:0-0(2000).
DR EMBL: AF139911; AAF44343.1; -.
SQ SEQUENCE 90 AA; 10368 MW; C3D573E78B52ABFA CRC64;

Query Match 74.1%; Score 43; DB 13; Length 90;
Best Local Similarity 77.8%; Pred. No. 0.69;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
DB 26 HWSYGLWPG 34

RESULT 6
O9KM11 PRELIMINARY; PRT; 565 AA.
AC O9KM11:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE HYPOTHETICAL PROTEIN VCA0578.
GN VCA0578.
OS Vibrionaceae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrion.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=204066833; PubMed=10952301;

RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
RT *cholerae*.";
RL Nature 406:477-483(2000).
DR EMBL: AE004388; AAF96480.1; -.
DR TIGR: VCA0578; -
SQ SEQUENCE 565 AA; 64837 MW; 36GD8479E2ADE303 CRC64;

Query Match 72.4%; Score 42; DB 2; Length 565;
Best Local Similarity 66.7%; Pred. No. 6.5;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
DB 179 HWRFYLRPG 187

RESULT 7
O82588 PRELIMINARY; PRT; 508 AA.
AC O82588:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE F1104.7 PROTEIN (FRAGMENT).
GN F1104.7.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;

RA WASHU;
RT "The A. thaliana Genome Sequencing Project";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Abu-Threideh J., Stoneking T., Langston Y., Trevaaskis E.,
RT "The sequence of A. thaliana F1104.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Waterston R.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF096370; AAC62783.1; -.
DR INTERPRO: IPR002885; -.
DR PFM; PF01535; PPR; 9.
FT NON_TER 508 508
SQ SEQUENCE 508 AA; 56879 MW; EB1449A49067AEDD CRC64;

Query Match 67.2%; Score 39; DB 10; Length 508;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 WSYXLRPG 10
DB 80 WCYSLRPG 87

RESULT 8
O9SZRS PRELIMINARY; PRT; 828 AA.
AC O9SZRS:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE HYPOTHETICAL 93.2 KDA PROTEIN.
GN F27B13.190 OR ATG629350.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Zimmermann W., Gruenisen A., Wambutt R., Bancroft I.,
RA Mewes H.W., Mayer K.F.X., Schueller C.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Zimmermann W., Gruenisen A., Wambutt R., Kalicki J., Wohlmann P.,
RA Smith A., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL050352; CAB43670.1; -.
DR EMBL: AL161575; CAB79753.1; -.
DR INTERPRO: IPR000195; -.
RN [5]
SQ SEQUENCE 828 AA; 93177 MW; CAB050A332BC5706 CRC64;

Query Match 67.2%; Score 39; DB 10; Length 828;
Best Local Similarity 66.7%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
 |||||
 Db 114 HWSYGWLP 122

RESULT 9
 O9W7G0 PRELIMINARY; PRT; 33 AA.
 AC O9W7G0;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN) (FRAGMENT).
 GN GNRH2.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99312119; PubMed=10385393;
 RA Von Schalburg K.R., Sherwood N.M.;
 RT "Regulation and expression of gonadotropin-releasing hormone gene differs in brain and gonads in rainbow trout.";
 RL Endocrinology 140:3012-3024(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Von Schalburg K.R., Sherwood N.M.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 DR EMBL: AF110993; ADD43463.1; -.
 DR INTERPRO: IPR002012; -.
 DR PFW: PFO0446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation.
 FT NON_TER 33
 FT SEQUENCE 33 AA; 3668 MW; 099C825E4A72A3BB CRC64;
 SO

Query Match 65.5%; Score 38; DB 13; Length 33;
 Best Local Similarity 66.7%; Pred. No. 2.1;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
 |||||
 Db 25 HWSYGWLP 33

RESULT 10
 O9PT34 PRELIMINARY; PRT; 33 AA.
 AC O9PT34;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE GONADOTROPIN-RELEASING HORMONE (FRAGMENT).
 GN GNRH1.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99312119; PubMed=10385393;
 RA Von Schalburg K.R., Sherwood N.M.;
 RT "Regulation and expression of gonadotropin-releasing hormone gene differs in brain and gonads in rainbow trout.";
 RL Endocrinology 140:3012-3024(1999).
 RN

RN [2]
 RP SEQUENCE FROM N.A.
 RA Von Schalburg K.R., Sherwood N.M.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 CC EMBL: AF110533; ADD43461.1; -.
 DR INTERPRO: IPR002012; -.
 DR PFW: PFO0446; GNRH; 1.
 DR PROSITE: PS00256; AKH; UNKNOWN_1.
 DR PROSITE: PS00473; GNRH; 1.
 FT NON_TER 33
 FT SEQUENCE 33 AA; 3741 MW; 1FE1535E742B7EBB CRC64;
 SO

Query Match 65.5%; Score 38; DB 13; Length 33;
 Best Local Similarity 66.7%; Pred. No. 2.1;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
 |||||
 Db 25 HWSYGWLP 33

RESULT 11
 O9Z094 PRELIMINARY; PRT; 82 AA.
 AC O9Z094;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN) PRECURSOR.
 GN PREPRO-GNRH-1.
 OS Oncorhynchus nerka (Sockeye salmon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8023;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=NIKKO; TISSUE=BRAIN;
 RX MEDLINE=96020547; PubMed=8546809;
 RA Ashihara M., Suzuki M., Kubokawa K., Yoshitani Y., Kobayashi M., Urano A., Aida K.;
 RT "Two differing precursor genes for the salmon-type gonadotropin-releasing hormone exist in salmonids.";
 RL J. Mol. Endocrinol. 15:1-9(1995).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 DR EMBL: D31868; BAA06666.1; -.
 DR INTERPRO: IPR002012; -.
 DR PFW: PFO0446; GNRH; 1.
 DR PROSITE: PS00256; AKH; UNKNOWN_1.
 DR PROSITE: PS00473; GNRH; 1.
 DR PRODOM: PD006549; -; 1.
 KW Signal; Hormone; Amidation.
 FT SIGNAL 1 23
 FT CHAIN 24 33
 FT CHAIN 37 82
 FT CHAIN 82 82
 FT SEQUENCE 82 AA; 9126 MW; C64044EA521B2B8B CRC64;
 SO

Query Match 65.5%; Score 38; DB 13; Length 82;
 Best Local Similarity 66.7%; Pred. No. 5.2;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
 |||||
 Db 25 HWSYGWLP 33

RESULT 12
 ID 09W7G1 PRELIMINARY: PRT: 82 AA.
 AC 09W7G1:
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE GONADOTROPIN-RELEASING HORMONE (GNRH) (LULIBERIN).
 GN GNRL.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99312119; PubMed=10385393;
 RT Von Schalburg K.R., Sherwood N.M.;
 RT "Regulation and expression of gonadotropin-releasing hormone gene
 RT differs in brain and gonads in rainbow trout.";
 RL Endocrinology 140:3012-3024(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Von Schalburg K.R., Sherwood N.M.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC FOLLICLE-STIMULATING HORMONES.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 DR EMBL; AF110992; AAD3462.1; -.
 DR INTERPRO: IPR002012; -.
 DR INTERPRO: IPR002047; -.
 DR PFAM: PF00446; GNRH; 1.
 DR PROSITE: PS00256; ARH; UNKNOWN_1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Hormone; Amidation.
 SQ SEQUENCE 82 AA; 9232 MW; 7595B4FCC65FDFD6 CRC64;

Query Match 65.5%; Score 38; DB 13; Length 82;
 Best Local Similarity 66.7%; Pred. No. 5.2;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 HWSYXLRPG 10

 DB 25 HWSYXLRPG 33

 RESULT 13
 ID 0918Q0 PRELIMINARY: PRT: 82 AA.
 AC 0918Q0:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE PRO-SGNRH-I.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE-BRAIN:
 RC Ferriere F., Bailhache T., Jégo P.;
 RT "Oncorhynchus mykiss sgnrh-I cDNA from brain."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF232212; AAF91280.1; -.
 SQ SEQUENCE 82 AA; 9198 MW; 7595A0B896556A69 CRC64;

Query Match 65.5%; Score 38; DB 13; Length 82;
 Best Local Similarity 66.7%; Pred. No. 5.2;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10

 DB 25 HWSYXLRPG 33

 RESULT 14
 ID 0918P9 PRELIMINARY: PRT: 82 AA.
 AC 0918P9:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE PRO-SGNRH-II.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN:
 RA Ferriere F., Bailhache T., Jégo P.;
 RT "Oncorhynchus mykiss sgnrh-II cDNA in the brain."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF232213; AAF91281.1; -.
 SQ SEQUENCE 82 AA; 9203 MW; 8053B4F221A0FF08 CRC64;

Query Match 65.5%; Score 38; DB 13; Length 82;
 Best Local Similarity 66.7%; Pred. No. 5.2;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 HWSYXLRPG 10

 DB 25 HWSYXLRPG 33

 RESULT 15
 ID 09PSY9 PRELIMINARY: PRT: 88 AA.
 AC 09PSY9:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE SGNRH (FRAGMENT).
 OS Sparus aurata (Gilthead sea bream).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
 OC Sparidae; Sparus.
 OX NCBI_TaxID=8175;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary:
 RA Nabissi M.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF046799; AAD02425.1; -.
 DR INTERPRO: IPR002012; -.
 DR PFAM: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 FT NON_TER 88
 SQ SEQUENCE 88 AA; 9788 MW; F7EB868C2FBD19F CRC64;

Query Match 65.5%; Score 38; DB 13; Length 88;
 Best Local Similarity 66.7%; Pred. No. 5.6;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 HWSYXLRPG 10

 DB 25 HWSYXLRPG 33

Fri Mar 2 12:19:26 2001

us-09-306-689-2.rspt

Search completed: March 2, 2001, 11:00:45
Job time: 455 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:51:20 ; Search time 56.87 Seconds
(without alignments)
16.715 Million cell updates/sec

Title: US-09-306-689-2
Perfect score: 58
Sequence: 1 XHMSYXLRPGXXXX 14

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	87.9	10	1	RHPGC
2	51	87.9	10	1	RHSHG
3	51	87.9	67	2	gonadoliberin - sh
4	51	87.9	89	2	gonadoliberin prec
5	51	87.9	90	1	RHMSG
6	51	87.9	92	1	RHHUG
7	51	87.9	92	1	RHRTG
8	47	81.0	10	1	RHAQ1
9	47	81.0	92	2	I50644
10	45	77.6	98	2	gonadotropin-relea
11	42	72.4	551	2	E64728
12	42	72.4	565	2	G82443
13	41	70.7	80	1	RHID1S
14	39	67.2	508	2	T01937
15	39	67.2	828	2	T08556
16	38	65.5	10	2	A21114
17	38	65.5	74	2	I51092
18	38	65.5	82	2	I51180
19	38	65.5	82	2	I51355
20	38	65.5	82	2	I51365
21	38	65.5	82	2	I51331
22	38	65.5	90	2	A23735
23	38	65.5	90	2	I51095
24	38	65.5	205	2	T36753
25	37	63.8	10	2	A49187
26	37	63.8	584	2	UC1229
27	36	62.1	954	2	S60301
28	35	60.3	82	2	G69446
29	35	60.3	161	2	C75438

30	35	60.3	220	2	T23358	hypothetical prote
31	35	60.3	390	2	S54026	ribosomal protein
32	35	60.3	477	1	C62LMA	cellulase (EC 3.2.
33	35	60.3	492	2	C83592	hypothetical prote
34	35	60.3	526	2	G72091	4-alpha-glucanotra
35	35	60.3	611	2	S57527	DNA helicase II -
36	35	60.3	691	2	D71430	hypothetical prote
37	35	60.3	711	1	A47136	macrophage-stimula
38	35	60.3	1232	2	T47993	hypothetical prote
39	34	58.6	119	2	F72531	hypothetical prote
40	34	58.6	138	2	G71506	ribosomal protein
41	34	58.6	316	2	A53440	aldose reductase h
42	34	58.6	351	2	T28864	hypothetical prote
43	34	58.6	486	2	S55146	probable ribosomal
44	34	58.6	676	2	A45900	complement C3b rec
45	34	58.6	972	2	T47922	anthranilate phosph

ALIGNMENTS

```

RESULT 1
RHPGC
gonadoliberin - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 18-Mar-1997
C:Accession: A01411
R:Baba, Y.; Matsuo, H.; Schally, A.V.
Biochem. Biophys. Res. Commun. 44, 459-463, 1971
A:Title: Structure of the porcine LH- and FSH-releasing hormone. II. Confirmation of
A:Reference number: A90172; MUID:72114303
A:Accession: A01411
A:Molecule type: protein
A:Residues: 1-10 <BAB>
R:Matsuo, H.; Arimura, A.; Nair, R.M.G.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 822-827, 1971
A:Title: Synthesis of the porcine LH- and FSH-releasing hormone by the solid-phase me
A:Reference number: A90176; MUID:72065376
A:Contents: annotation; synthesis
A:Note: the synthetic and natural hormones have the same physicochemical and biologic
R:Baba, Y.; Arimura, A.; Schally, A.V.
Biochem. Biophys. Res. Commun. 45, 483-487, 1971
A:Title: On the tryptophan residue in porcine LH and FSH-releasing hormone.
A:Reference number: A90175; MUID:72117544
A:Contents: annotation
A:Note: Trp-3 appears to be essential for biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match      87.9%  Score 51:  DB 1:  Length 10:
Best Local Similarity 88.9%  Pred. No. 0.0019:
Matches      8:  Conservative      0:  Mismatches      1:  Indels      0:  Gaps      0:

OY      2  HMSYXLRPG 10
DB      2  HMSYXLRPG 10

RESULT 2
RSHG
gonadoliberin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Mar-1997
C:Accession: A93780; A01411
R:Butcher, R.; Butcher, M.; Amos, M.; Ling, N.; Monahan, M.; Rivier, J.; Fellows, R.;
Proc. Natl. Acad. Sci. U.S.A. 69, 278-282, 1972
A:Title: Primary structure of the ovine hypothalamic luteinizing hormone-releasing fa
A:Reference number: A93780; MUID:72094314
A:Accession: A93780

```

A:Molecule type: protein
A:Residues: 1-10

A>Note: the natural and synthetic hormones have the same biological activity
C:Comment: This hypothalamic hormone stimulates the secretion of both luteinizing and follicle stimulating hormone
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.0019;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
|||||
DB 2 HWSYGLRPG 10

RESULT 3
178541
gonadoliberin precursor - rhesus macaque (fragment)
N:Alternate names: luteinizing hormone releasing hormone
C:Species: Macaca mulatta (Rhesus macaque)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C:Accession: I78541
R:Ma, Y.J.; Costa, M.E.; Ojeda, S.R.
Neuroendocrinology 60, 346-359, 1994
A:Title: Developmental expression of the genes encoding transforming growth factor alpha
A:Reference number: I58134; M0ID:95124501
A:Status: preliminary;
A:Molecule type: mRNA
A:Residues: 1-67 <RES>
A:Cross-references: GB:S75918; NID:g912831; PIDN:AA33096.1; PID:g912832
C:Superfamily: gonadoliberin

Query Match 87.9%; Score 51; DB 2; Length 67;
Best Local Similarity 88.9%; Pred. No. 0.012;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
|||||
DB 7 HWSYGLRPG 15

RESULT 4
151423
gonadoliberin precursor - African clawed frog
N:Alternate names: luteinizing hormone releasing hormone
C:Species: Xenopus laevis (African clawed frog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: I51423
R:Hayes, W.P.; Wray, S.; Battey, J.F.
Endocrinology 134, 1835-1845, 1994
A:Title: The frog GnRH-I gene has a mammalian-like expression pattern and conserved domain
A:Reference number: I51423; M0ID:94185563
A:Accession: I51423
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-89 <HAY>
A:Cross-references: GB:L28040; NID:g496291; PIDN:AAA9728.1; PID:g496292
C:Genetics:
A:Gene: GnRH-I
C:Superfamily: gonadoliberin

Query Match 87.9%; Score 51; DB 2; Length 89;
Best Local Similarity 88.9%; Pred. No. 0.016;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10

DB 25 HWSYGLRPG 33
|||||

RESULT 5
RHMSG
gonadoliberin precursor - mouse
N:Alternate names: gonadotropin-releasing hormone (GnRH); luteinizing hormone releasing hormone (LHRH); gonadoliberin; gonadoliberin-associated protein (GAP)
C:Species: Mus musculus (house mouse)
C>Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
C:Accession: A47578
R:Wason, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikoli
Science 234, 1366-1371, 1986
A:Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible for the hypogonadism in the mouse
A:Reference number: A47578; M0ID:87069928
A:Accession: A47578
A:Molecule type: DNA
A:Residues: 1-90 <MAS>
A:Cross-references: EMBL:M14872; NID:g193576; PIDN:AAA37717.1; PID:g387175
C:Genetics:
A:Introns: 45/3; 77/3
C:Function:
A:Description: gonadoliberin stimulates pituitary secretion of luteinizing hormone and follicle stimulating hormone
A>Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:22-31/Product: gonadoliberin #status predicted <GLR>
F:35-90/Product: gonadoliberin-associated protein (Gln) (in mature form) #status predicted
F:22/Modified site: pyroglutamate carboxylic acid (Gln) (in mature form) #status predicted
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form) from following

Query Match 87.9%; Score 51; DB 1; Length 90;
Best Local Similarity 88.9%; Pred. No. 0.016;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
|||||
DB 23 HWSYGLRPG 31

RESULT 6
RHMSG
gonadoliberin precursor - human
N:Alternate names: gonadotropin releasing hormone (GnRH); luteinizing hormone releasing hormone (LHRH); gonadoliberin-associated protein (GAP); progadoliberin
C:Species: Homo sapiens (man)
C>Date: 17-Mar-1987 #sequence_revision 21-Jul-1995 #text_change 18-Jun-1999
C:Accession: S05308; A26173; A93342; A90108; A01410; S45718
R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
Nucleic Acids Res. 17, 6403-6404, 1989
A:Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone
A:Reference number: S05308; M0ID:8936682
A:Accession: S05308
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-92 <HAY>
A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956
R:Adelman, J.P.; Wason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone
A:Reference number: A94090; M0ID:86094338
A:Accession: A26173
A:Molecule type: mRNA
A:Residues: 1-92 <ADE>
A:Cross-references: GB:M12578; NID:g183418; PIDN:AAA35916.1; PID:g386749
R:Seeburg, P.H.; Adelman, J.P.
Nature 311, 666-668, 1984
A:Title: Characterization of cDNA for precursor of human luteinizing hormone releasing hormone
A:Reference number: A93342; M0ID:85012739

A:Accession: A93342
A:Molecule type: mRNA
A:Residues: 1-15, 'S', 17-92 <SEB>
A:Cross-references: GB:X01059; NID:g34356; PIDN:CAA25526.1; PID:g34357
A:Experimental source: placenta
R:Tan, L.; Rousseau, P.
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982
A:Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in
A:Reference number: A90108; MUID:83126573
A:Accession: A90108
A:Molecule type: protein
A:Residues: 24-33 <TAN>
A:Experimental source: placental trophoblasts
R:Leibovitz, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amsterde
FEBS Lett. 346, 203-206, 1994
A:Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by th
A:Reference number: S45718; MUID:94283597
A:Contents: annotation; degradation pathway of synthetic hormone
C:Genetics:
A:Gene: GDB:GNRH; LHRH; GRH
A:Cross-references: GDB:133746; OMIM:227200; OMIM:152760
A:Map position: bp21-8p11.2
A:Introns: 47/3; 79/3
C:Function:
A:Description: gonadoliberin stimulates pituitary secretion of luteotropin and follitropin
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-92/Product: progonaoliberin #status predicted <PCN>
F:24-33/Product: gonadoliberin #status experimental <MAT>
F:37-92/Product: gonadoliberin-associated protein #status predicted <GAP>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

Query Match 87.9%; Score 51; DB 1; Length 92;
Best Local Similarity 88.9%; Pred. No. 0.017;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 HWSYLRPG 10
Db 25 HWSYLRPG 33

RESULT 7
RHRTS
gonadoliberin precursor - rat
N:Alternate names: gonadoliberin-associated protein (GAP); gonadotropin releasing hormo
N:Contents: gonadoliberin; prolactin release-inhibiting factor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
C:Accession: A40147; B26173; A48410
R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
Mol. Endocrinol. 3, 1257-1262, 1989
A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic ex
A:Reference number: A40147; MUID:89384661
A:Accession: A40147
A:Molecule type: DNA
A:Residues: 1-92 <BON>
A:Cross-references: GB:M1670; NID:g204447; PIDN:AAA41264.1; PID:g204448
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A:Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadot
A:Reference number: A94090; MUID:86094338
A:Accession: B26173
A:Molecule type: mRNA
A:Residues: 1-92 <ADP>
A:Cross-references: GB:M12579; NID:g204445; PIDN:AAA41263.1; PID:g204446
R:Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blalock, J.E.
Cell. Mol. Neurobiol. 12, 447-454, 1992
A:Title: Thymocytes express a mRNA that is identical to hypothalamic luteinizing hormone
A:Reference number: A48410; MUID:93105480

A:Accession: A48410
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <MAI>
A:Cross-references: GB:S50870; NID:g262059; PIDN:AAB24572.1; PID:g262060
A:Experimental source: thymus
A:Note: sequence extracted from NCBI backbone (NCBI:121082, NCBI:P.121083)
C:Genetics:
A:Introns: 47/3; 79/3
C:Function:
A:Description: stimulates pituitary secretion of luteotropin and follitropin
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activi
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-92/Product: progonaoliberin #status predicted <PCN>
F:24-33/Product: gonadoliberin #status predicted <GIN>
F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 87.9%; Score 51; DB 1; Length 92;
Best Local Similarity 88.9%; Pred. No. 0.017;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 HWSYLRPG 10
Db 25 HWSYLRPG 33

RESULT 8
RHAQI
gonadoliberin I - American alligator
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Alligator mississippiensis (American alligator)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Mar-1997
C:Accession: A60066
R:Lovejoy, D.A.; Fischer, W.H.; Parker, D.B.; McRory, J.E.; Park, M.; Lance, V.; Swan
Regul. Pept. 33, 105-116, 1991
A:Title: Primary structure of two forms of gonadotropin-releasing hormone from brains
A:Reference number: A60066; MUID:91352338
A:Accession: A60066
A:Molecule type: protein
A:Residues: 1-10 <LOV>
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:10/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 81.0%; Score 47; DB 1; Length 10;
Best Local Similarity 77.8%; Pred. No. 0.01;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 HWSYLRPG 10
Db 2 HWSYLRPG 10

RESULT 9
I50644
gonadoliberin I precursor - chicken
N:Alternate names: gonadotropin-releasing hormone I
C:Species: Gallus gallus (chicken)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
C:Accession: I50644; S33507
R:Dunn, I.C.; Chen, Y.; Hook, C.; Sharp, P.J.; Sang, H.M.
J. Mol. Endocrinol. 11, 19-29, 1993
A:Title: Characterization of the chicken pregonadotropin-releasing hormone-I gene
A:Reference number: I50644; MUID:94059355
A:Accession: I50644
A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-92 <DD2>
A:Cross-references: EMBL:X69491; NID:g496326; PIDN:CAA9246.1; PID:g311612
C:Genetics: 47/3; 79/3
A:Introns: 47/3; 79/3
C:Superfamily: gonadoliberin

Query Match 81.0%; Score 47; DB 2; Length 92;
Best Local Similarity 77.8%; Pred. No. 0.089;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
11111111
Db 25 HWSYXLRPG 33

RESULT 10
150739
gonadotropin-releasing hormone - Cichlid (Haplochromis burtoni)
C:Species: Haplochromis burtoni
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: 150739
R:White, S.A.; Kasten, T.L.; Bond, C.T.; Adelman, J.P.; Fernald, R.D.
Proc. Natl. Acad. Sci. U.S.A. 92, 8363-8367, 1995
A:Title: Three gonadotropin-releasing hormone genes in one organism suggest novel roles
A:Reference number: 150739; MUID:95396797
A:Accession: 150739
A:Status: preliminary; translated from GH/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-98 <WHI>
A:Cross-references: EMBL:U31865; NID:g905398; PIDN:MAC59691.1; PID:g905399
C:Superfamily: gonadoliberin

Query Match 77.6%; Score 45; DB 2; Length 98;
Best Local Similarity 77.8%; Pred. No. 0.22;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
11111111
Db 24 HWSYXLRPG 32

RESULT 11
E64728
yabn protein - Escherichia coli
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 08-Oct-1999
C:Accession: E64728
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.
A: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:9742617
A:Accession: E64728
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-551 <BIAT>
A:Cross-references: GB:AE000117; GB:U00096; NID:g1786250; PIDN:AACT3180.1; PID:g1786256;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: yabn

Query Match 72.4%; Score 42; DB 2; Length 551;
Best Local Similarity 66.7%; Pred. No. 4.2;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
11111111
Db 179 HWSYXLRPG 187

RESULT 12
G82443
conserved hypothetical protein VCA0578 [imported] - Vibrio cholerae (group O1 strain
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: G82443
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.
charadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: G82443
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-565 <HEI>
A:Cross-references: GB:AE004388; GB:AE003853; NID:g9657979; PIDN:AAF96480.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene position: 2

Query Match 72.4%; Score 42; DB 2; Length 565;
Best Local Similarity 66.7%; Pred. No. 4.3;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
11111111
Db 179 HWSYXLRPG 187

RESULT 13
RH1D5
gonadoliberin I precursor - Sharptooth catfish
N:Alternate names: gonadoliberin, catfish-type; gonadotropin-releasing hormone I (GHR
N:Contents: gonadoliberin I; gonadoliberin I-associated protein form I; gonadoliberin
C:Species: Clarias gariepinus (Sharptooth catfish)
C>Date: 30-Sep-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
C:Accession: S45602; S45601; JCI1242; S42936; S42937
R:Bogerd, J.; Zandbergen, T.; Andersson, E.; Goos, H.
Eur. J. Biochem. 222, 541-549, 1994
A:Title: Isolation, characterization and expression of cDNAs encoding the catfish-type
A:Reference number: S45600; MUID:94291651
A:Accession: S45602
A:Molecule type: mRNA
A:Residues: 1-80 <BOG1>
A:Cross-references: EMBL:X78049; NID:g459433; PIDN:CAA54971.1; PID:g459434
A:Note: gonadoliberin I-associated protein form I
A:Accession: S45601
A:Molecule type: protein
A:Residues: 1-46, 'S', 48-59, 'G', 61-80 <BOG2>
A:Cross-references: EMBL:X78048; NID:g459431; PIDN:CAA54970.1; PID:g459432
A:Note: gonadoliberin I-associated protein form II, presumed to be a polymorphic form
R:Bogerd, J.; Li, K.W.; Janssen-Dommerholt, C.; Goos, H.
Biochem. Biophys. Res. Commun. 187, 127-134, 1992
A:Title: Two gonadotropin-releasing hormones from African catfish (Clarias gariepinus
A:Reference number: JCI1242; MUID:92392313
A:Accession: JCI1242
A:Molecule type: protein
A:Residues: 22-31 <BOG3>
A:Experimental source: brain
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1/21/Domain: signal sequence #status predicted <SIG>
F:22-31/Product: gonadoliberin I #status experimental <MAT1>
F:33-80/Product: gonadoliberin I-associated protein #status predicted <MAT2>
F:22/Modified site: pyroglutamate carboxylic acid (Glu) (in mature form) #status experi
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match 70.7%; Score 41; DB 1; Length 80;

Best Local Similarity 66.7%; Pred. No. 0.95;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
|||:| ||
Db 23 HWSHGLNPG 31

RESULT 14

T01937

hypothetical protein F1104.7 - Arabidopsis thaliana (fragment)

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 24-Mar-1999

C:Accession: T01937

R:Abu-Threideh, J.; Stoneking, T.; Langston, Y.; Trevaaskis, E.

submitted to the EMBL Data library, October 1998

A:Description: The sequence of A. thaliana F1104.

A:Reference number: 214466

A:Accession: T01937

A:Status: translated from GR/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1508 <ABU>

A:Cross-references: EMBL:AF096370; NID:g3695372; PID:g3695380

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 4

A:Note: F1104.7

Query Match

67.2%; Score 39; DB 2; Length 508;

Best Local Similarity 75.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 WSYXLRPG 10
| | | |||
Db 80 WCYSLRPG 87

RESULT 15

T08556

hypothetical protein F27B13.190 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999

C:Accession: T08556

R:Bevan, M.; Zimmermann, W.; Gruenelsen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; May

submitted to the Protein Sequence Database, May 1999

A:Reference number: 216442

A:Accession: T08556

A:Molecule type: DNA

A:Residues: 1828 <BEV>

A:Cross-references: EMBL:AL050352; GSPDB:GN00062; ATSP:F27B13.190

A:Experimental source: cultivar Columbia; BAC clone F27B13

C:Genetics:

A:Gene: ATSP:F27B13.190

A:Map position: 4

A:Introns: 52/2; 87/1; 146/1; 777/3

Query Match

67.2%; Score 39; DB 2; Length 828;

Best Local Similarity 66.7%; Pred. No. 22;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
|||:| ||
Db 114 HWSYFOAPG 122

Search completed: March 2, 2001, 10:55:38
Job time: 258 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:54:05 ; Search time 32.03 Seconds

(without alignments)
14.115 Million cell updates/sec

Title: US-09-306-689-2
Perfect score: 58
Sequence: 1 XHMSYXLRPEXXXX 14

Scoring table:

BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwisProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	51	87.9	61	GONI_SHEEP
2	51	87.9	63	GONI_MESAU
3	51	87.9	67	GONI_MACMU
4	51	87.9	89	GONI_XENIA
5	51	87.9	90	GONI_MOUSE
6	51	87.9	91	GONI_PIG
7	51	87.9	92	GONI_HUMAN
8	51	87.9	92	GONI_RAT
9	51	87.9	92	GONI_TUPGB
10	47	81.0	10	GONI_ALAMI
11	47	81.0	92	GONI_CHICK
12	45	77.6	94	GONI_HAPBU
13	45	77.6	95	GONI_PAGMA
14	45	77.6	95	GONI_SPAU
15	42	72.4	92	GONI_CAVRO
16	42	72.4	551	GONI_CAVRO
17	41	70.7	80	GONI_ECOLI
18	38	65.5	10	GONI_CLATA
19	38	65.5	14	GONI_ONCKE
20	38	65.5	74	GONI_ONCKE
21	38	65.5	82	GONI_ONCKE
22	38	65.5	82	GONI_SALSA
23	38	65.5	82	GONI_SALSA
24	38	65.5	89	GONI_SALSA
25	38	65.5	90	GONI_HAPBU
26	38	65.5	90	GONI_PAGMA
27	38	65.5	90	GONI_SPAU
28	38	65.5	94	GONI_CAVRO
29	38	65.5	94	GONI_CAVRO
30	37	63.8	10	GONI_PETMA
31	37	63.8	584	GONI_PETMA
32	35	60.3	390	GONI_PETMA
33	35	60.3	477	GONI_PETMA

ALIGNMENTS

RESULT ID	1	STANDARD:	PRT:	61 AA.
GONI_SHEEP	1			
GONI_SHEEP	1			
AC	Q28588:			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	PROGONADOLIBERIN I PRECURSOR (CONTRAINS: GONADOLIBERIN I (LHRH I) (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I) (FRAGMENT).			
DE	GNRH1 OR GNRH OR LHRH.			
OS	Ovis aries (Sheep).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Caprinae; Ovis.			
RN	[1]			
RP	SEQUENCE OF 12-61 FROM N.A.			
RC	STRAIN-WESTERN RANGE: TISSUE-HYPOTHALAMUS;			
RA	Rodriguez R.E., Wise M.E.;			
RL	Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE OF 1-10.			
RA	MEDLINE-72094314; PubMed-4550508;			
RA	Burgos R., Butcher M., Amoss M., Ling N., Monahan M., Rivier J.,			
RA	Fellows R., Blackwell R., Vale W., Guillemin R.;			
RT	"Primary structure of the ovine hypothalamic luteinizing hormone-			
RT	releasing factor (LRF) (LH-hypothalamus-LRF-gas chromatography-mass			
RT	spectrometry-decapeptide-Edman degradation).";			
RL	Proc. Natl. Acad. Sci. U.S.A. 69: 278-282(1972).			
CC	- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS: IT STIMULATES			
CC	THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING			
CC	HORMONES.			
CC	- SIMILARITY: BELONGS TO THE GNRH FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL; U02517; AA03433.1; -.			
DR	PIR; A93780; RSHSG.			
DR	INTERPRO; IPR002012; -.			
DR	PFAM; PF00446; GNRH; 1.			
DR	PROSITE; PS00473; GNRH; 1.			
KW	Cleavage on pair of basic residues; Hormone; Amideation; Hypothalamus;			
KW	Placenta.			
FT	NON_TER	1		
FT	CHAIN	1	>61	PROGONADOLIBERIN I.
FT	PEPTIDE	1	10	GONADOLIBERIN I.
FT	PEPTIDE	14	>61	GNRH-ASSOCIATED PEPTIDE I.
FT	ACT_SITE	3	3	APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT				ACTIVITY.

FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP).
FT NON_TER 61 61
SQ SEQUENCE 61 AA; 6828 MW; 63962A1AE319B8F0 CRC64;

Query Match 87.9%; Score 51; DB 1; Length 61;
Best Local Similarity 88.9%; Pred. No. 0.0045;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
Db 2 HWSYGLRPG 10

RESULT 2

CON1_MESAU STANDARD; PRT; 63 AA.
AC 009163;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
(LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]
(FRAGMENT).
DE GNRH1 OR GNRH OR LHRH.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
RN [1]
RP SEQUENCE FROM N.A.
RA Jansen H.T., Stevens P.J., Zeidler P., Lehman M.N.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. IT STIMULATES
THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
HORMONES.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U91938; AAB51302.1; -.
DR INTERPRO: IPR002012; -.
DR PFAM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta.
KW NON_TER 1 1
FT CHAIN 1 >63 PROGNADOLIBERIN I.
FT PEPTIDE 1 10 GONADOLIBERIN I.
FT PEPTIDE 14 >63 GNRH-ASSOCIATED PEPTIDE I (BY
SIMILARITY).
FT ACT_SITE 3 3 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
ACTIVITY (BY SIMILARITY).
FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID (BY
SIMILARITY).
FT MOD_RES 10 10 AMIDATION (G-11 PROVIDE AMIDE GROUP) (BY
SIMILARITY).
FT NON_TER 63 63
SQ SEQUENCE 63 AA; 7370 MW; FC94995676F77180 CRC64;

Query Match 87.9%; Score 51; DB 1; Length 63;
Best Local Similarity 88.9%; Pred. No. 0.0046;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
Db 2 HWSYGLRPG 10

RESULT 3

CON1_MACMU STANDARD; PRT; 67 AA.
AC P55247;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
(LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I]
(FRAGMENT).
DE GNRH1 OR GNRH OR LHRH.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-HYPOTHALAMUS;
RX MEDLINE-95124501; Pubmed=7545971;
RA Ma Y.J., Costa M.E., Ojeda S.R.;
RT "Developmental expression of the genes encoding transforming growth
factor alpha and its receptor in the hypothalamus of female rhesus
macaques";
RL Neuroendocrinology 60:346-359(1994).
CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. IT STIMULATES
THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
HORMONES.
CC -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S75918; AAB33096.1; -.
DR INTERPRO: IPR002012; -.
DR PFAM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
KW NON_TER 1 1
FT SIGNAL <1 5 BY SIMILARITY.
FT CHAIN 6 >67 PROGNADOLIBERIN I.
FT PEPTIDE 6 15 GONADOLIBERIN I.
FT PEPTIDE 19 >67 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 8 8 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
ACTIVITY (BY SIMILARITY).
FT MOD_RES 6 6 PYROLIDONE CARBOXYLIC ACID (BY
SIMILARITY).
FT MOD_RES 15 15 AMIDATION (G-16 PROVIDE AMIDE GROUP) (BY
SIMILARITY).
FT NON_TER 67 67
SQ SEQUENCE 67 AA; 7573 MW; 505394DAA261A3F2 CRC64;

Query Match 87.9%; Score 51; DB 1; Length 67;
Best Local Similarity 88.9%; Pred. No. 0.0049;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
Db 7 HWSYGLRPG 15


```

RESULT 4
GONL_XENLA STANDARD: PRT: 89 AA.
AC P45656:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE (LH-RH) (LULIBERIN I).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FOREBRAIN;
RX MEDLINE=9418563; PubMed=8137750;
RA Hayes W.P., Wray S., Battey J.F.;
RT "The frog gonadotropin-releasing hormone-I (GNRH-I) gene has a
RT mammalian-like expression pattern and conserved domains in
RT GNRH-associated peptide, but brain onset is delayed until
RT metamorphosis."
RL Endocrinology 134:1835-1844(1994).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L28040; AAA49728.1; -
DR INTERPRO; IPR002012; -
DR PFM; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal.
FT SIGNAL 1 23
FT CHAIN 24 89
FT PEPTIDE 24 33
FT CHAIN 37 89
FT PEPTIDE 37 85
FT MOD_RES 24 24
FT MOD_RES 33 33
FT SEQUENCE 89 AA; 10246 MW; 6FAF36FBAE0D4284 CRC64;

Query Match 87.9%; Score 51; DB 1; Length 89;
Best Local Similarity 88.9%; Pred. No. 0.0065;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYLRPG 10
DB 25 HWSYLRPG 33

RESULT 5
GONL_MOUSE STANDARD: PRT: 90 AA.
AC P13562:
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR (CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR
DE I).
GN GNRH1 OR GNRH.
OS Mus musculus (Mouse).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87069928; PubMed=3024317;
RA Mason A.J., Haylick J.S., Zoeller R.T., Young W.S. III,
RA Phillips H.S., Nikolics K., Seeburg P.H.;
RT "A deletion truncating the gonadotropin-releasing hormone gene is
RT responsible for hypogonadism in the hpg mouse."
RL Science 234:1366-1371(1986).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -----
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14872; AAA37717.1; -
DR MGD; MGI:95789; GNRH.
DR INTERPRO; IPR002012; -
DR PFM; PF00446; GNRH; 1.
DR PROSITE; PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 21
FT CHAIN 22 90
FT PEPTIDE 22 31
FT PEPTIDE 35 90
FT ACT_SITE 24 24
FT MOD_RES 22 22
FT MOD_RES 31 31
FT SEQUENCE 90 AA; 10337 MW; 1C0766F8A8264D9 CRC64;

Query Match 87.9%; Score 51; DB 1; Length 90;
Best Local Similarity 88.9%; Pred. No. 0.0065;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYLRPG 10
DB 23 HWSYLRPG 31

RESULT 6
GONL_PIG STANDARD: PRT: 91 AA.
AC P49921:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR (CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I).
GN GNRH1 OR GNRH.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HYPOPHALAMUS;
RA Weesner G.D., Mattern R.L., Becker B.A.;
RL Submitted (May-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 24-33.
RX MEDLINE=72114303; PubMed=4946067;
RA Baba Y., Matsuo H., Schally A.V.;

```

RT "Structure of the porcine LH- and FSH-releasing hormone. II.
RT Confirmation of the proposed structure by conventional sequential
RT analyses.";
RL Biochem. Biophys. Res. Commun. 44:459-463(1971).
RN [3]
RP SYNTHESIS OF GONADOLIBERIN.
RX MEDLINE=72065376; PubMed=4942726;
RA Matsuo H., Arimura A., Nair R.M.G., Schally A.V.;
RT "Synthesis of the porcine LH- and FSH-releasing hormone by the solid-
RT phase method.";
RL Biochem. Biophys. Res. Commun. 45:822-827(1971).
RN [4]
RP SYNTHESIS OF GONADOLIBERIN.
RX MEDLINE=72117544; PubMed=4946275;
RA Baba Y., Arimura A., Schally A.V.;
RT "On the tryptophan residue in porcine LH and FSH-releasing hormone.";
RL Biochem. Biophys. Res. Commun. 45:483-487(1971).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L32864; AAA31066.1; -;
CC PIR: A01411; RHHGG.
CC DR INTERPRO: IPR002012; -;
CC PRAM: PF00446; GNRH: 1.
CC DR PROSITE: PS00473; GNRH: 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC Placenta; Signal.
CC KW CHAIN 1 23
CC FT CHAIN 1 23 PROGONADOLIBERIN I.
CC FT PEPTIDE 24 33 GONADOLIBERIN I.
CC FT PEPTIDE 34 91 GNRH-ASSOCIATED PEPTIDE I.
CC FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
CC ACTIVITY.
CC FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
CC FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
CC SO SEQUENCE 91 AA; 10090 MW; 8340474F32DDAA99 CRC64;
CC -----
CC Query Match 87.9%; Score 51; DB 1; Length 91;
CC Best Local Similarity 88.9%; Pred. No. 0.0066;
CC Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 HWSYXLRPG 10
DB 25 HWSYXLRPG 33
CC -----
CC RESULT 7
CC GON1_HUMAN STANDARD; PRT; 92 AA.
CC AC P01148;
CC DT 21-JUL-1986 (Rel. 01, Created)
CC DT 01-APR-1988 (Rel. 07, Last sequence update)
CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
CC DE PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
CC (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
CC HORMONE I) (GNRH I) (LULIBERIN I) (GONADORELIN); GNRH-ASSOCIATED
CC PEPTIDE I].
CC DE GNRH1 OR GNRH OR LHRH.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CC [1];

RP SEQUENCE FROM N.A.
RX MEDLINE=8936682; PubMed=2671939;
RA Hayflick J.S., Adelman J.P., Seeburg P.H.;
RT "The complete nucleotide sequence of the human gonadotropin-releasing
RT hormone gene.";
RL Nucleic Acids Res. 17:6403-6403(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094338; PubMed=2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=85012739; PubMed=6090951;
RA Seeburg P.H., Adelman J.P.;
RT "Characterization of cDNA for precursor of human luteinizing hormone
RT releasing hormone.";
RL Nature 311:666-668(1984).
RN [4]
RP SEQUENCE OF 24-33.
RX MEDLINE=83126573; PubMed=6760865;
RA Tan L., Rousseau P.;
RT "The chemical identity of the immunoreactive LHRH-like peptide
RT biosynthesized in the human placenta.";
RL Biochem. Biophys. Res. Commun. 109:1061-1071(1982).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- PHARMACEUTICAL: AVAILABLE UNDER THE NAMES FACTREL (AYERST LABS),
CC LUTERPUSE OR LUTERLEF (FERRING PHARMACEUTICALS) AND RELISORM
CC (SERONO).
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X01059; CAA25526.1; -;
CC DR EMBL: M12578; AAA35916.1; -;
CC DR EMBL: X15215; CAA33285.1; -;
CC DR PIR: A01410; RHHGG.
CC DR PIR: A26173; A26173.
CC DR PIR: S05308; S05308.
CC DR MIM: 152760; -;
CC DR INTERPRO: IPR002012; -;
CC PRAM: PF00446; GNRH: 1.
CC DR PROSITE: PS00473; GNRH: 1.
CC Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
CC KW Placenta; Pharmaceutical; Signal.
CC KW CHAIN 1 23
CC FT CHAIN 1 23 PROGONADOLIBERIN I.
CC FT PEPTIDE 24 33 GONADOLIBERIN I.
CC FT PEPTIDE 37 92 GNRH-ASSOCIATED PEPTIDE I.
CC FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
CC ACTIVITY.
CC FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
CC FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
CC FT CONFLICT 16 16 W -> S (TN REF. 3).
CC SO SEQUENCE 92 AA; 10380 MW; 30A72221B076FA79 CRC64;
CC -----
CC Query Match 87.9%; Score 51; DB 1; Length 92;
CC Best Local Similarity 88.9%; Pred. No. 0.0067;
CC Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 HWSYXLRPG 10

|||||
25 HWSYGLRPG 33

RESULT 8
GONI_RAT STANDARD: PRT: 92 AA.
AC P07490;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE LUTEINIZING HORMONE RELEASING HORMONE I] (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR
CN GNRH1 OR GNRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-86094338; PubMed-2867548;
RA Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;
RT "Isolation of the gene and hypothalamic cDNA for the common precursor
RT of gonadotropin-releasing hormone and prolactin release-inhibiting
RT factor in human and rat."
RL Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-89384661; PubMed-2476669;
RA Bond C.T., Hayflick J.S., Seeburg P.H., Adelman J.P.;
RT "The rat gonadotropin-releasing hormone: SH locus: structure and
RT hypothalamic expression."
RL Mol. Endocrinol. 3:1257-1262(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-THYMUS;
RX MEDLINE-93105480; PubMed-1468115;
RA Walter C.C., Marchetti B., Lebeuf R.D., Blalock J.E.;
RT "Thymocytes express a mRNA that is identical to hypothalamic
RT luteinizing hormone-releasing hormone mRNA."
RL Cell. Mol. Neurobiol. 12:447-454(1992).
RN [4]
RP SEQUENCE OF 1-47 FROM N.A.
RC TISSUE-HEART;
RX MEDLINE-87149087; PubMed-3547652;
RA Adelman J.P., Bond C.T., Douglass J., Herbert E.;
RT "Two mammalian genes transcribed from opposite strands of the same
RT DNA locus."
RL Science 235:1514-1517(1987).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: S50870; AAB24572.1; -
DR EMBL: M12579; AAA41263.1; -
DR EMBL: M31670; AAA41264.1; -
DR EMBL: M15527; AAA42141.1; ALT_SEQ.
DR EMBL: M15529; AAA42139.1; -
DR EMBL: M15528; -; NOT_ANNOTATED_CDS.
DR PIR: B26173; RHRTG.
DR PIR: A48410; A48410.
DR INTERPRO: IPR002012; -

DR PFAM: PF00446; GNRH: 1.
DR PROSITE: PS00473; GNRH: 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92 PROGNADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 92 PROLACTIN RELEASE-INHIBITING FACTOR I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT MOD_RES 24 24 ACTIVITY.
FT MOD_RES 33 33 PYRROLIDONE CARBOXYLIC ACID.
FT SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64; AMIDATION (G-34 PROVIDE AMIDE GROUP).
Query Match 87.9%; Score 51; DB 1; Length 92;
Best Local Similarity 88.9%; Pred. No. 0.0067;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 2 HWSYGLRPG 10
DB 25 HWSYGLRPG 33
RESULT 9
GONI_TUPGB STANDARD: PRT: 92 AA.
AC Q95335;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGNADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE LUTEINIZING HORMONE RELEASING HORMONE I] (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
CN GNRH1 OR GNRH.
OS Tupia glis belangeri (Common tree shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupaiia.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-HYPOTHALAMUS;
RX MEDLINE-97079639; PubMed-8921350;
RA Kassten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,
RA Fernald R.D.;
RT "Characterization of two new preproGNRH mRNAs in the tree shrew:
RT first direct evidence for mesencephalic GNRH gene expression in a
RT placental mammal."
RL Gen. Comp. Endocrinol. 104:7-19(1996).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U63326; AAB16837.1; -
DR INTERPRO: IPR002012; -
DR PFAM: PF00446; GNRH: 1.
DR PROSITE: PS00473; GNRH: 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92 BY SIMILARITY.
FT PEPTIDE 24 33 PROGNADOLIBERIN I.
FT PEPTIDE 37 92 GONADOLIBERIN I.
FT ACT_SITE 26 26 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACTIVITY.

FT	MOD_RES	24	24	PYRROLIDONE CARBOXYLIC ACID (BY SIMILARITY)
FT	MOD_RES	33	33	AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY SIMILARITY)
FT	SEQUENCE	92 AA; 10197 MW; 4FDBF2C58CF5F63B CRC64;		
Query Match				
Best Local Similarity	87.9%;	Score 51; DB 1; Length 92;		
Matches	8; Conservative	0; Mismatches	1; Indels	0; Gaps
QY	2 HWSYXLRPG 10			
Db	25 HWSYGLRPG 33			
RESULT 10				
ID	GONI_ALIMI	STANDARD;	PRT;	10 AA.
AC	P37041; P20407;			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	GONADOLIBERIN I (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I) (LH-RH I) (LULIBERIN I).			
OS	Alligator mississippiensis (American alligator).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Archosauria; Crocodylidae; Alligatorinae; Alligator.			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE-BRAIN.			
RA	Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,			
RA	Lance V., Swanson P., Rivier J.E., Sherwood N.M.;			
RT	"Primary structure of two forms of gonadotrophin-releasing hormone from brains of the American alligator (Alligator mississippiensis).";			
RL	Regul. Pept. 33:105-116(1981).			
CC	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.			
CC	-1- SIMILARITY: BELONGS TO THE GNRH FAMILY.			
DR	PIR; A60066; RHAQ1.			
DR	INTERPRO: IPR002012; -			
DR	PFAM: PF00446; GNRH; 1.			
DR	PROSITE: PS00473; GNRH; 1.			
KW	Hormone; Amidation; Hypothalamus.			
FT	MOD_RES	1		PYRROLIDONE CARBOXYLIC ACID.
FT	MOD_RES	10		AMIDATION.
SO	SEQUENCE	10 AA; 1172 MW; 284B23D7286B45A3 CRC64;		
Query Match				
Best Local Similarity	81.0%;	Score 47; DB 1; Length 10;		
Matches	7; Conservative	1; Mismatches	1; Indels	0; Gaps
QY	2 HWSYXLRPG 10			
Db	2 HWSYGLRPG 10			
RESULT 11				
ID	GONI_CHICK	STANDARD;	PRT;	92 AA.
AC	P37042; P20407;			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-JUN-1984 (Rel. 29, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	PROGONADOLIBERIN I PRECURSOR (CONTAINS: GONADOLIBERIN I (LHRH I) (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING HORMONE I) (GNRH I) (LULIBERIN I) (GNRH-ASSOCIATED PEPTIDE I).			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
RN	[1].			

```

RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEGHORN:
RX MEDLINE=94059355; PubMed=7902095;
RA Dunn I.C., Chen Y., Hook C., Sharp P.J., Sang H.M.;
RT "Characterization of the chicken preprogonadotrophin-releasing
RL hormone-I gene.";
RN J. Mol. Endocrinol. 11:19-29(1993).
RP
RC SEQUENCE OF 24-33.
RX TISSUE=HYPOPHALAMUS;
RA MEDLINE=82265778; PubMed=7050119;
RA King J.A., Miller R.P.;
RT "Structure of chicken hypothalamic luteinizing hormone-releasing
RL hormone. II. Isolation and characterization.";
RN J. Biol. Chem. 257:10729-10732(1982).
RP
RC [3]
RX SEQUENCE OF 24-33.
RA TISSUE=HYPOPHALAMUS;
RC King J.A., Miller R.P.;
RT "Structure of avian hypothalamic gonadotrophin-releasing hormone.";
RL S. Afr. J. Sci. 78:124-125(1982).
RP
RC [4]
RX SYNTHESIS OF 24-33.
RA MEDLINE=82265777; PubMed=7050118;
RA King J.A., Miller R.P.;
RT "Structure of chicken hypothalamic luteinizing hormone-releasing
RL hormone. I. Structural determination on partially purified
RT material.";
RL J. Biol. Chem. 257:10722-10728(1982).
RP
RC [5]
RX FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
RL -I- SIMILARITY: BELONGS TO THE GnRH FAMILY.
RP
RC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69491; CAA49246.1; -.
DR PIR; S33507; S33507.
DR INTERPRO: IPR002012; -.
DR PPM; PP00446; GnRH; 1.
DR PROSITE; PS00473; GnRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW signal.
FT FT SIGNAL. 1 23
FT CHAIN 24 92 PROGONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 92 GnRH-ASSOCIATED PEPTIDE I.
FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT MOD_RES 33 33
SQ SEQUENCE 92 AA; 10206 MW; 61AE87EBAF508B6A CRC64.

```

DE (LH-RH I) (LULIBERIN I).
 GN GNRH1.
 OS Haplochromis burtoni.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
 CC Cichlidae; Astrotellipia.
 [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=95396797; PubMed=7667296;
 RA White S.A., Kasten T.L., Bond C.T., Adelman J.P., Fernald R.D.;
 RT "Three gonadotropin-releasing hormone genes in one organism suggest
 RT novel roles for an ancient peptide."
 RL Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).
 [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=99061842; PubMed=9843638;
 RA White R.B., Fernald R.D.;
 RT "Ontogeny of gonadotropin-releasing hormone (GnRH) gene expression
 RT reveals a distinct origin for GnRH-containing neurons in the
 RT midbrain."
 RL Gen. Comp. Endocrinol. 112:322-329(1998).
 [3]
 RN SEQUENCE OF 23-32.
 RP TISSUE=PIUITARY;
 RC MEDLINE=95372591; PubMed=7644702;
 RA Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,
 RA White S.A., Francis R.C., Fernald R.D., Licht P., Warty C.,
 RA Sherwood N.M.;
 RT "Primary structure of solitary form of gonadotropin-releasing hormone
 RT (GnRH) in cichlid pituitary: three forms of GnRH in brain of cichlid
 RT and pumpkinseed fish."
 RL Regul. Pept. 57:43-53(1995).
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE
 CC RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALMIC-PITUITARY-
 CC GONADAL AXIS.
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS
 CC TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPOTHYSEAL AXONS.
 CC -1- MASS SPECTROMETRY: MW=1113.9; METHOD=MALDI; RANGE=23-32.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U31865; AAC59691.1; -;
 DR EMBL: AF076961; AAC27716.1; -;
 DR INTERPRO: IPR002012; -;
 DR PFAM: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 RN Signal: Multigene family.
 FT SIGNAL 1 22
 FT CHAIN 1 22 PROGONADOLIBERIN I.
 FT PEPTIDE 23 32 GONADOLIBERIN I.
 FT MOD_RES 36 94 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
 FT MOD_RES 23 23 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 32 32 AMIDATION (G-33 PROVIDE AMIDE GROUP).
 FT CONFLICT 86 94 ENGRTEPK -> KMDCHSNERFL (IN REF. 1).
 SQ SEQUENCE 94 AA; 10382 MW; E57DBA8333278D7 CRC64;

RESULT 13
 GONL_PAGMA STANDARD; PRT; 95 AA.
 AC P70074;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
 DE (LH-RH I) (LULIBERIN I).
 OS Pagrus major (Red sea bream) (Chrysophrys major).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 CC Sparidae; Chrysophrys.
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=BRAIN;
 RC Okuzawa K., Granneman J., Bogerd J., Goos H., Zohar Y., Kagawa H.;
 RL Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: D86582; BAA13129.1; -;
 DR INTERPRO: IPR002012; -;
 DR PFAM: PF00446; GNRH; 1.
 DR PROSITE: PS00473; GNRH; 1.
 KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
 RN Signal: Multigene family.
 FT SIGNAL 1 23
 FT CHAIN 1 23 POTENTIAL.
 FT PEPTIDE 24 95 GONADOLIBERIN I.
 FT MOD_RES 37 95 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
 FT MOD_RES 24 24 PYRROLIDONE CARBOXYLIC ACID (BY
 FT SIMILARITY).
 FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP)
 FT (BY SIMILARITY).
 SQ SEQUENCE 95 AA; 10566 MW; 61E79C990328D73E CRC64;

Query Match 77.6%; Score 45; DB 1; Length 95;
 Best Local Similarity 77.8%; Pred. No. 0.085;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 HWSYLRPG 10
 |||||
 DB 25 HWSYLRPG 33

RESULT 14
 GONL_SPAU STANDARD; PRT; 95 AA.
 AC P51919;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
 DE (LH-RH I) (LULIBERIN I) (SBGNH).
 OS Sparus aurata (Gilthead sea bream).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 CC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
 CC Sparidae; Sparus.

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA MEDLINE=95268499; PubMed=7749463;
RX Gottlieb Y., Elizur A., Chow M., Chen T.T., Zohar Y.,
RT "Molecular cloning and characterization of a novel gonadotropin-
RL releasing hormone from the gilthead seabream (Sparus aurata).";
RM Mol. Mar. Biol. Biotechnol. 4:27-35(1995).
RN [2]
RP SEQUENCE OF 26-35.
RC TISSUE-BRAIN;
RA MEDLINE=95083645; PubMed=7991588;
RX Powell J.F.F., Zohar Y., Elizur A., Park M., Fischer W.H.,
RA Craig A.G., Rivier J.E., Lovejoy D.A., Sherwood N.M.;
RT "Three forms of gonadotropin-releasing hormone characterized from
RL brains of one species.";
RM Proc. Natl. Acad. Sci. U.S.A. 91:12081-12085(1994).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC -1- MASS SPECTROMETRY: MW=1113.6; METHOD=MALDI; RANGE=26-35.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U30320; AAA75469.1; -
DR INTERPRO: IPR002012; -
DR PFAM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Signal; Multigene family.
FT SIGNAL 1 25
FT CHAIN 26 95 PROGONADOLIBERIN I.
FT PEPTIDE 26 95 GONADOLIBERIN I.
FT MOD_RES 39 95 GNRH-ASSOCIATED PEPTIDE I (POTENTIAL).
FT MOD_RES 26 26 PYROLIDONE CARBOXYLIC ACID.
FT MOD_RES 35 35 AMIDATION (G-36 PROVIDE AMIDE GROUP).
SQ SEQUENCE 95 AA; 10753 MW; 49313FD6D6B87DA CRC64;

Query Match 77.6%; Score 45; DB 1; Length 95;
Best Local Similarity 77.8%; Pred. No. 0.085;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
Db 27 HWSYGLSPG 35

RESULT 15
ID GON1_CAVPO STANDARD; PRT; 92 AA.
AC 054713;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I)
DE (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING
DE HORMONE I) (GNRH I) (LULIBERIN I); GNRH-ASSOCIATED PEPTIDE I].
GN GNRH1 OR GNRH OR LHRH.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HARTLEY WHITE; TISSUE=HYPOPHALAMUS;
RX MEDLINE=97462693; PubMed=9322920;
RA Jimenez-Linan M., Rubin B.S., King J.C.;
RT "Examination of guinea pig luteinizing hormone-releasing hormone gene

```

```

RT reveals a unique decapeptide and existence of two transcripts in the
RT brain.";
RL Endocrinology 138:4123-4130(1997).
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES
CC THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING
CC HORMONES.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF033346; AAB87688.1; -
DR INTERPRO: IPR002012; -
DR PROSITE: PS00473; GNRH; 1.
KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.
FT SIGNAL 1 23
FT CHAIN 24 92 POTENTIAL.
FT PEPTIDE 24 92 PROGONADOLIBERIN I.
FT PEPTIDE 24 92 GONADOLIBERIN I.
FT ACT_SITE 26 26 GNRH-ASSOCIATED PEPTIDE I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
FT ACT_SITE 26 26 ACTIVITY (BY SIMILARITY).
FT MOD_RES 24 24 PYROLIDONE CARBOXYLIC ACID (BY
FT MOD_RES 24 24 SIMILARITY).
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP) (BY
FT MOD_RES 33 33 SIMILARITY).
SQ SEQUENCE 92 AA; 10279 MW; ACF74613F456D663 CRC64;

Query Match 72.4%; Score 42; DB 1; Length 92;
Best Local Similarity 66.7%; Pred. No. 0.29;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
Db 25 YWSYGVRRG 33

```

Search completed: March 2, 2001, 11:01:19
Job time: 434 sec

FILE REFERENCE: PC10202A
CURRENT APPLICATION NUMBER: US/09/506,078
CURRENT FILING DATE: 2000-02-16
EARLIER APPLICATION NUMBER: N/A
EARLIER FILING DATE: 1999-02-17
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 27
LENGTH: 442
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
US-09-506-078-27
4GnRH-LmGD-4GnRH encoded by POE-GnRH:gd:GnRH

alignment_scores:

Quality: 304.50 Length: 634
Ratio: 1.208 Gaps: 25
Percent Similarity: 39.748 Percent Identity: 22.082

alignment_block:

US-09-306-689-12 x US-09-506-078-27 ..

Align seg 1/1 to: US-09-506-078-27 from: 1 to: 442

```
178 CAGATGAGACTAGCGCTCGCCCTGGCAGCGGTCTCAAGATTGAG 227
    ::::::::::::::::::::
19  GluHisTrpSerTyrglyLeuArgProly.....GluHisTrpSe 32
228 CTAGCGCTCGCTCGCGGTGAGCTAGCCAGCATGAGCTAGCGCTGC 277
    |||||||||||||||
32  rTyrglyLeuArgProgly.....GluHisTrpSerTyrglyLeuA 46
278 GCCCTGGCAGCGGTAGCCAGATTGAGCTAGCGCTCGCGGTGGA 327
    |||||||
46  rGProgly.....GluHisTrpSerTyrglyLeuArgProglySer 59
328 TCTAGCTTCCCAAAACCTGGGCAAAAAATTTCCCTATATCC... 375
    |||||
60  MetSerLeuPro...ThrProAlaProArgValThrValTyValAspPr 75
376 .....CAAATTACCAATATGATGATGCAACAGATAATG 409
    ::|||
75  oProAlaTyrProMetProArgTyraSerTy...ThrLuarGtrpHisT 91
410 GTTACAGAGATTAGTCAAAAGCG.....GCCGAAGAGTTGGGAT 450
    ::|
91  hrThnglyProIleProSerProPheAlaAspGlyArgGluGlnProVal 107
451 GAGGTACAAAGAGAAGAACGCAATATATTGCACACAGCTCAACACAGTT 500
    |||||
108 GluValArg..... 110
501 AGGCAGCATCAAAACCGCTATTGGCTTAACAGCGTGAGCTGTGTAT 550
    |||
111 .....TyrlaThrSerAlaAlaAlaCysAspMetLeuAlaLeuIleA 125
551 CCGCTCCCAAAATGATTAATTTGCTACGAAACAAATAAGAGCCAGCA 600
    ::|||
125 laAspProGlnVal.....GlyArgThr 132
601 TTAGTCTGCGCGAAAGCATGTGTACAAATGCAATAAAGCCAAACGT 650
    |||
133 LeuTrpGlnAlaValArgArgHisAlaArgAlaTyraSerAlaThrValI 149
651 ATTATCTGCGCATCAATCTATTGAGCTCAGTATTGGCTGGAATGAT 700
    |||
149 eTrpTyrlaSerIleGluSerGlyCysAlaArgProLeuSerTyrlaMetGluT 166
701 TAGATGAGCGCTTACAGATAACAGCAACCAACATGCTTGTGTAACAGT 750
    |||
```

```
166 yTrhnglyCysGluProArgLysHisPheGlyTyrcysArgTyArgThr 182
751 GGCTTGAGCTAACAAATTCATTATGAAATATTCCTATTCAGAAA 800
    ::|
183 Pro...ProPheTrpAspSerPheLeuAlaGlyPheAla....Tyrrp 196
801 AACACTGAGCAATTTGGTGGAGCAAAATTAATGCAATTTGTTCAAAACTAC 850
    |||
196 oThrAspAspGlnLeuGly..... 202
851 AAAATATCAAAAGCTTAGGCACTTAGAGACAAACATCAAAATATCGGT 900
202 ..... 202
901 GCACTGTGTAAGCTGGCCTTGTGTTAGATGATATCTCAGGCTATTATC 950
202 ..... 202
951 GGGCGCAACAGCTGCACTTGACTTGCGAGATAAATGCTTCAACAGCTA 1000
    |||
203 .....LeuIleMet.....AlaAlaProAla 210
1001 AAAAAGTGGGTGGCGTTTGAATTTGCCAAACCAAGTTGTTGTAATATT 1050
    ::|
210 rgleuValAlaGluGlyGlnTyraArgAlaLeuTyrlaAspGly..... 224
1051 ACCAAGCGCTTCTCTCATTTTAGCCCAACGCTTGACAGAGCTTT 1100
    |||
225 ThrValAlaTyrlaTrpAspPheMetVal..... 233
1101 ATCTCAACTGGCGCTGCTGCTTAATGCTTCTACTGTTCTCTTG 1150
233 ..... 233
1151 CGATTAGCCCATTTAGCATTTGCCGTAATGCCGTAATTAATCATGCA 1200
233 ..... 233
1201 AAAAGTTAGAGATTATGCCGACGCTTAAATAATTAGCTATGACGG 1250
    |||
234 ...SerLeuProAlaGlyAspCysTyrrPheSerLysLeuGly..... 246
1251 AGATTAATTTATACAGATAATACAGCGGGACAGGACATATGATGCAT 1300
246 ..... 246
1301 CGTTTACTGCATTAATACCGCATTTGCCGCTATTGCTGGTGTGCTCT 1350
    |||
247 .....AlaAlaArgGlyTyrrPheGlyAlaCys 256
1351 GCTGCTGCGCGGATTTAACATTTGAAAGGTTAAACATATATCTGTGCAT 1400
    |||
257 PheProAlaArgAspTyrglyGlnGlnLysValLeuArgLeuTrpTyLe 273
1401 CACGAAAT.....AGCAAAAAAGAAAGTGACCATTCAAACCTGT 1441
    |||
273 uThrGlnTyrrProGlnGlnAlaHisLysAlaIleValAspTyrrTrp 290
1442 TCCGAGAG.....GCTGATTTTGTCTAAAGAACTCCTAAT 1476
    |||
290 hemeTarGHisGlyGlyValValProProTyrrPheGlnGlnSerLysGly 306
1477 TATAAGCAACTAA.....GA 1493
307 TyrglnProProProAlaAlaAspGlyLysSerProAlaProProGlyAs 323
1494 TGAGAAATCGAAGAAATCATCGCTCAAAATGGCAGCGGATCACTCAA 1543
    |||
323 pAspGlnAlaArgGlnAspGlnGlyGlnThrGlnAspGlyValAlaIleGlyA 340
1544 AGCAAGTTGATGATCTTATCGCAAAAGGTAAAGGCAAAATTAACCAAGAT 1593
    ::|
340 rGlnu.....GlyAsnGlyGlyProProGlyPro 349
```



```

1594 GACGATCAAAAGTTGTGATTAATGATGATGCTCAACATAGCAAAA 1643
    |||
350 Gln..... 350
1644 TGTGACAAACAGCTTAGATTAATTCATCTGTAAGTCATTACCT 1693
    |||
351 .....GlyaspGlyuserGlnThrProGlnAlaAsnGly 363
1694 CGCTATATGTTGAGAAATGATTAAGTGGCTCCAACTTCAATGTTGGAT 1743
    |||
363 LysAlaGlnGlyProGlyProGlyProSerProAspAlaAspArgPro 379
1744 CAAGTTTATCTCTCTCAATTT..... 1767
    |||
380 GlnGlyThrProSerLeuGlnAlaIleThrHisProProAlaProAl 396
1768 .....GCTAGGGATCTCAGCATTCAGCTAGCGCTGCGCCGTGCA 1810
    |||
396 arThrProAlaArgAlaProGlnHisThrSerThrGlyLeuArgProGly 412
1811 GCGGTTCTCAAGTTGAGCTAGCGCTGCGCTGCGGTGCTAGCCAG 1860
    |||
413 .....Gln 413
1861 CATTTGAGCTAGCGCTGCGCTGCGCGAGCGGTAGCCAAAGATTGAGCTA 1910
    |||
414 HisThrSerThrGlyLeuArgProGly.....GlnHisThrSerThr 427
1911 CGGCTGCGCTGCGGTGATTCAGCATTCAGCTAGCGCTGCGCCGTG 1960
    |||
427 rGlyLeuArgProGly.....GlnHisThrSerThrGlyLeuArgProG 442
1961 GC 1962
    ||
442 Ty 442

```

seq_name: /cgn2_6/ptodata/2/paa/US60_COMB.pep:US-60-120-454-27

```

seq_documentation_block:
; Sequence 27, Application US/60120454
; GENERAL INFORMATION:
; APPLICANT: Campos, Manuel
; APPLICANT: Yule, Terecila D
; APPLICANT: Matindog, Serge
; APPLICANT: Dutsch, Becky A
; TITLE OF INVENTION: DUAL IMMUNE RESPONSE
; FILE REFERENCE: PCI10202
; CURRENT APPLICATION NUMBER: US/60-120-454
; EARLIER APPLICATION NUMBER: N/A
; EARLIER FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 27
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: 46NH-tmgd-46NH encoded by pQE-gnRH:gd:gnRH
US-60-120-454-27

```

```

alignment_scores:
    Quality: 304.50      Length: 634
    Ratio: 1.208        Gaps: 25
    Percent Similarity: 39.748    Percent Identity: 22.082

```

alignment_block:
US-09-306-689-12 x US-60-120-454-27 ..

```

Align seg 1/1 to: US-60-120-454-27 from: 1 to: 442
178 CAGCATTTGAGCTAGCGCTGCGCCCTGCGAGCGGTTCTCAAGATTGAG 227
    |||
19 GlnHisThrSerThrGlyLeuArgProGly.....GlnHisThrSer 32
228 CTAGGGCTGCGCGGTGAGCTAGCGCTGCGCGGTGAGCTAGCGCTGCG 277
    |||
32 rGlyLeuArgProGly.....GlnHisThrSerThrGlyLeuArg 46
278 GCGCTGCGAGCGGTAGCCAAAGATTGAGCTAGCGCTGCGCGGTGAG 327
    |||
46 rProGly.....GlnHisThrSerThrGlyLeuArgProGlySer 59
328 TCTAGCTTCCCAAACTGCGGCAAAAATTATCTGTAATTTCC... 375
    |||
60 MetSerLeuPro...ThrProAlaProArgValThrValValAspArg 75
376 .....CAAAATTACCAATATGATTAAGTGAACAAGTAATG 409
    |||
75 oProAlaThrProMetProArgThrAsnThr...ThrGlnArgThrHis 91
410 GTTTACAGGATTTAGTCAAGCG.....GCCGAAAGTTGGGATTT 450
    |||
91 hrThrGlyProIleProSerProPhaAlaAspGlyArgGlnGlnProVal 107
451 GAGGTACAAAGAGAGACGACCAATTAATTCAGACAGCTCAAAACAGTTT 500
    |||
108 GlnValArg..... 110
501 AGGCAAGATTCAAAACCGCTATTGGCTTAAGTGAAGCTGAGCTATGTT 550
    |||
111 .....TyraThrSerAlaAlaIleCysAspMetLeuAlaLeuIle 125
551 CCGCTCCACAAATTGATAATTGCTACAGAAAACCTAAACAGGCCAAGCA 600
    |||
125 LaspProGlnVal.....GlyArgThr 132
601 TTAGGTTCTGCCGAAGCATTTGTACAAATGCAAAATTAAGCCAAACCTG 650
    |||
133 LeuThrGlnAlaValArgArgHisAlaArgAlaValArgAlaValIle 149
651 ATTATGTCATTCATTAATTTAGCTCAGATTTGGCTGGAATGGAT 700
    |||
149 eThrPylrLysIleGlySerGlyCysAlaArgProLeuThrYrMerGln 166
701 TAGATGAGGCTTACAGATAACAGCAACCAACATGCTTCTCTAAAGCT 750
    |||
166 yThrGlnGlySerProAlaGlyHisPhaGlyThrCysArgGlyArgThr 182
751 GCGCTGAGCTAACAAATTCATTAATTTGAATAATTTGCTAATTCAGTAA 800
    |||
183 Pro...ProThrPaspSerPheLeuAlaGlyPheAla.....Tyrr 196
801 AACACTTGACGAATTTGCTGACCAAAATTTAGTCATTTGTTGTTCAAA 850
    |||
196 oThrAspAspGlnLeuGly..... 202
851 AAAATATCAAGGCTTAGGACTTTAGAGACAACTCAAAATATTCGCT 900
    |||
202 ..... 202
901 GGACTTGATTAAGCTGCGCTGCTTGTAGATGTTATCTCAGGCTATTATC 950
    |||
202 ..... 202
951 GGGCGCAACAGCTGCATCTGTACTTGACAGATAAAATGTTCAACAGCTA 1000
    |||
203 .....LeuIleMet.....AlaAlaProAla 210
1001 AAAAAGTGGTGCGGGTTTGAATTTGCAAAACCAAGTTGTTGTAATATT 1050
    |||
210 rGlnValGlnGlyGlnThrArgArgAlaLeuThrIleAspGly..... 224

```

```

1051 ACCAAGCCGTTCTCTTACATTTTACCCCAACGTTGTCAGCAGGTTT 1100
    ||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
225 ThrValAlaTyrThrAspPheMetVal..... 233
1101 ATCTTCAACTGGCCTGTGGCTTATTCCTTACTCTTCTCTCTG 1150
233 ..... 233
1151 CGATTAGCCCATTTAGCATTTGCGGTATTGCGGATAATTAATCATCA 1200
233 ..... 233
1201 AAAAGTTAGAGATTATGCCGAACGCTTTAAAAATTAGCCTATGACGG 1250
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
234 ...SerIeuProAlaGlyAspCysTyrPheSerIeuGly..... 246
1251 AGATTAATTATTACGACAAATATCAGCGGGGACAGGACTATTGATGAT 1300
246 ..... 246
1301 CGGTACTGCAATTAAATACGCGCTTGGCGCTATTGCTGTGTGTCTCT 1350
247 ..... 247
1351 GCTGTCGAGCCGATTAAACATTTGAAAAAGTTAAACATTAATCTGTGAT 1400
    ||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
257 PheProAlaIaArgAspTyrGluGlnLysValLeuArgLeuThrTyrLe 273
1401 CACGAAAT.....AGCAAAAAAGAGAAAGTGCATTTCAAAACTGCT 1441
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
273 uThGlnIuTyrTyrProGlnGlnAlaHisLysAlaIleValAspTyrTrp 290
1442 TCCGAGAG.....GCTGATTTTGGTAAAGAGTGGCTTAT 1476
    || : : : : : : : : : : : : : : : : : : : : : : : :
290 hemeTArgHisGlyGlyValValProProTyrPheGlnGlnSerLysGly 306
1477 TATAAGCAACTAAA..... 1493
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
307 TyrGluProProProAlaAlaAspGlyLysSerProAlaProProGlyAs 323
1494 TGAGAAATCAAGAAATCATCGGTCAAAATGGCGAGGATCACTCA 1543
    : : : : : : : : : : : : : : : : : : : : : : : :
323 pAspGlnAlaArgGlnAspGlnGlyGluThrGlnAspGlyAlaAlaGly 340
1544 AGCAAGTTGATGATCTTATCGCAAAAGSTAAGCGCAAAATTAACCAAGAT 1593
    : : : : : : : : : : : : : : : : : : : : : : : :
340 rGgLu.....GlyAsnGlyGlyProProGlyPro 349
1594 GAGCTATCAAAAGTTGTGATTAAGTAATGCTCAAAACATAGCAAAA 1643
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
350 Glu..... 350
1644 TGTGACAAACAGCTTAGTAATGTAATCTCATCTGTAAGTCATTTACT 1693
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
351 .....GlyAspGlyGlnSerGlnThrProGlnAlaAsnGly 363
1694 CGTCTAATGATTCGAGAAATGATATTAGTGCCTCAACTCAATCTTGAT 1743
    : : : : : : : : : : : : : : : : : : : : : : : :
363 lYAlaGlnGlyGluProLysProGlyProSerProAlaAspArgPro 379
1744 CAAAGTTATCTCTCTCAATT..... 1767
    : : : : : : : : : : : : : : : : : : : : : : : :
380 GluGlyTrpProSerLeuGlnAlaIleThrHisProProProAlaProAl 396
1768 .....GCTAGGGGATCTCAGCATGTGAGCTAGGCGCTGGCCCTGGCA 1810
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
396 aThProAlaArgAlaProGlnHisTyrSerTyrGlyLeuArgProGly 412
1811 GCGGTTTCAGATTTGAGCTACGGCCCTGCGGGGCTGAGCTAGCCAG 1860
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
113 .....Glu 413

```

```

1861 CATTGAGACTACGGCTGCGCCCTGGCAGCGGTAGCCAGATTGGACCTA 1910
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
414 HisTrpSerTyrGlyLeuArgProGly.....GlnHisTrpSerTyr 427
1911 CGGCTGCGTCCGGGTGATCTCAGCATTTGAGCTAGGCGCTGGCCCTG 1960
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
427 rGlyLeuArgProGly.....GlnHisTrpSerTyrGlyLeuArgProG 442
1961 GC 1962
    ||
442 Ly 442

seq_name: /cgn2_6/ptodata/2/paa/US090_COMB.pep:US-09-019-010-4

seq_documentation_block:
; Sequence 4, Application US/09019010
; GENERAL INFORMATION:
; APPLICANT: HARLAND, RICHARD
; APPLICANT: MANN, JOHN G.
; APPLICANT: ACRE, STEPHEN D.
; TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/019,010
; FILING DATE: 05-FEB-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,883
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0035
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 325-7812
; TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-019-010-4

alignment_scores:
    Quality: 290.00      Length: 49
    Ratio: 5.918        Gaps: 0
    Percent Similarity: 100.000    Percent Identity: 100.000

alignment_block:
US-09-306-689-12 x US-09-019-010-4 ..

Align seg 1/1 to: US-09-019-010-4 from: 1 to: 49

25 CAGCATTTGAGACTACGGCTGCGCCCTGGCAGCGGTCTCAAGATTGGAG 74
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlySerGlnAspTrpSe 17
75 CTACGGCTGCGTCCGGGTGCTCTAGCCAGCATTTGAGCTACGGCCCTGC 124

```

```

|||||
17 rtyrglyleuargproglyserglnhistrpserlytyglyleua 34
125 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCCTCGCGGT 171
34 rproglyserglyserglnasptpserlytyglyleuargprogly 49

```

seq_name: /cgn2_6/ptodata/2/paa/US092_COMB.pep:US-09-249-447A-7

seq_documentation_block:

```

; Sequence 7, Application US/09249447A
; GENERAL INFORMATION:
; APPLICANT: Manns, Jack G.
; TITLE OF INVENTION: PASSIVE IMMUNIZATION AS A TREATMENT FOR HORMONE
; FILE REFERENCE: 9001-0045
; CURRENT APPLICATION NUMBER: US/09/249,447A
; PRIOR FILING DATE: 1999-02-12
; PRIOR FILING DATE: 1998-02-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GnRH-2, Fig.
; US-09-249-447A-7

```

alignment_scores:

```

Quality: 290.00      Length: 49
Ratio: 5.918         Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

```

alignment_block:

US-09-306-689-12 x US-09-249-447A-7 ..

Align seg 1/1 to: US-09-249-447A-7 from: 1 to: 49

```

25 CAGCATTTGAGCTAGCGGCTCGCGCTGAGCGGTTCTCAAGATTGAG 74
|||||
1 Glnhistrpserlytyglyleuargproglyserglyserglnasptpse 17
75 CTACGGCTGCGGTCCGGGTGGCTCTAGCCAGCATTTGAGCTACGGCTGC 124
|||||
17 rtyrglyleuargproglyserglnhistrpserlytyglyleua 34
125 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCCTCGCGGT 171
|||||
34 rproglyserglyserglnasptpserlytyglyleuargprogly 49

```

seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:US-09-305-924-11

seq_documentation_block:

```

; Sequence 11, Application US/09305924A
; GENERAL INFORMATION:
; APPLICANT: Jack G. Manns
; APPLICANT: Stephen D. Acres
; APPLICANT: Richard Hatland
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305,924A
; PRIOR FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,217
; EARLIER FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 49
; TYPE: PRT

```

```

; ORGANISM: GnRH
US-09-305-924-11

```

alignment_scores:

```

Quality: 290.00      Length: 49
Ratio: 5.918         Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

```

alignment_block:

US-09-306-689-12 x US-09-305-924-11 ..

Align seg 1/1 to: US-09-305-924-11 from: 1 to: 49

```

25 CAGCATTTGAGCTAGCGGCTCGCGCTGAGCGGTTCTCAAGATTGAG 74
|||||
1 Glnhistrpserlytyglyleuargproglyserglyserglnasptpse 17
75 CTACGGCTGCGGTCCGGGTGGCTCTAGCCAGCATTTGAGCTACGGCTGC 124
|||||
17 rtyrglyleuargproglyserglnhistrpserlytyglyleua 34
125 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCCTCGCGGT 171
|||||
34 rproglyserglyserglnasptpserlytyglyleuargprogly 49

```

seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:US-09-306-689-11

seq_documentation_block:

```

; Sequence 11, Application US/09306689B
; GENERAL INFORMATION:
; APPLICANT: Robbins, Sarah C.
; TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
; FILE REFERENCE: 9001-0047
; CURRENT APPLICATION NUMBER: US/09/306,689B
; PRIOR FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: US 60/088,024
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; US-09-306-689-11

```

alignment_scores:

```

Quality: 290.00      Length: 49
Ratio: 5.918         Gaps: 0
Percent Similarity: 100.000  Percent Identity: 100.000

```

alignment_block:

US-09-306-689-12 x US-09-306-689-11 ..

Align seg 1/1 to: US-09-306-689-11 from: 1 to: 49

```

25 CAGCATTTGAGCTAGCGGCTCGCGCTGAGCGGTTCTCAAGATTGAG 74
|||||
1 Glnhistrpserlytyglyleuargproglyserglyserglnasptpse 17
75 CTACGGCTGCGGTCCGGGTGGCTCTAGCCAGCATTTGAGCTACGGCTGC 124
|||||
17 rtyrglyleuargproglyserglnhistrpserlytyglyleua 34
125 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCCTCGCGGT 171
|||||
34 rproglyserglyserglnasptpserlytyglyleuargprogly 49

```

```

seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:us-09-383-912-4
seq_documentation_block:
; Sequence 4, Application US/09383912
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNIS, JOHN G.
; TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016,22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-383-912-4

alignment_scores:
Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-306-689-12 x US-09-383-912-4 ..

Align seg 1/1 to: US-09-383-912-4 from: 1 to: 49
25 CAGCATTTGAGACTACGGCTCGGCTGAGCGAGGTTCTCAAGATTGGAG 74
|||||
1 GlnIstPrSerTyrGlyLeuAlaGlyProGlySerGlnIstPrSer 17
|||||
75 CTACGGCTCGGCTCGGCTGAGCTTACGACGATTTGAGTACGGCTGC 124
|||||
17 rTtGclYleuAlaGlyProGlySerGlnIstPrSerTyrGlyLeuAla 34
|||||
125 GCGTCCGAGCGGTAGCAAGATTTGAGTACGGCTCGGCTCGGCTG 171
|||||
34 rGPrGclYSerGlnIstPrSerTyrGlyLeuAlaGlyProGly 49

seq_name: /cgn2_6/ptodata/2/paa/US086_COMB.pep:us-08-669-785-2
seq_documentation_block:
; Sequence 2, Application US/08669785
; GENERAL INFORMATION:
; APPLICANT: Betsou, Fotini
; APPLICANT: Sebo, Peter

```

```

; APPLICANT: Guiso, Nicole
; TITLE OF INVENTION: Protective Epitopes Of Adenyli
; TITLE OF INVENTION: Cyclase-haemolysin(AC-Hly), Their Application To
; TITLE OF INVENTION: The Treatment Of To The Prevention Of Bordetella Infections
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,785
; FILING DATE: 27-JUN-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356,0072-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1706 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-669-785-2

alignment_scores:
Quality: 212.00 Length: 512
Ratio: 0.763 Gaps: 21
Percent Similarity: 54.297 Percent Identity: 22.266

alignment_block:
US-09-306-689-12 x US-08-669-785-2 ..

Align seg 1/1 to: US-08-669-785-2 from: 1 to: 1706
265 AGCTACGCGCTGCGCCTGCGAGCGGTAGCCAGATTGAGCTACGGCCT 314
|||||
349 AlaTyrGlyAla...AlaGlyLysSerLeuPheAspArgLysLeuGlyAl 364
|||||
315 GCGTCCGAGGTGATCTAGCTTCCCAAAACTGGGCAAAAATTTATCC 364
|||||
364 aAlaPrGlyAlaPrOserGlyArgSerLysPheSerProAspValLeuG 381
|||||
365 TCTATATTCGCCAAATTTACCAATATGATAGTACGAACAAGTAATGTTT 414
|||||
381 LuTrValPrOAlaSer.....ProGlyLeu 389
|||||
415 CAGCATTTAGTCAAGGCGCGAGAGTGGGATTTAGGTACAAAGAGA 464
|||||
390 Arg.....ArgProSerLeuGly...AlaValGlyAlaGly 400
|||||
465 AGAAGC.....ATAATATTGCAACAGCTCAAAAGCTT 499
|||||
400 nAspSerGlyTyrAspSerLeuAspGlyValGlySerArgSerPheSerL 417
|||||
500 TAGCAGAGATTCAACCGCTATTGGCTTAAGTACGAGCGCATGTGTTA 549
|||||
417 euGlyGlyValSerAspMet.....AlaIleVal 426
|||||
550 TCCGCTCCACAATTTGATTAATTGCTACAGAAAACCTAAGCAGGCCAAGC 599

```

[illegible][illegible]

seq_name: /cgn2_6/ptodata/2/paa/US095_COMB.pep:US-09-506-078-23

```
seq_documentation_block:
```

; GENERAL INFORMATION:

; TITLE OF INVENTION: FUSION PROTEINS

FILE REFERENCE: PC10202A

CURRENT FILING DATE: 2000-02-16

EARLIER APPLICATION NUMBER: N/A
FILING DATE: 1999-02-17

```

; NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 23
LENGTH: 417

```

```

;
; TYPE: PRT
OBCANTSM: Artificial Sequence

```

FEATURE: DESCRIBES THE
FEATURE INTRODUCED BY
; ;

OTHER INFORMATION: encoded by pQE-

alignment_scores:

Ratio: 4.088

[illegible]

US-09-306-689-12 x US-09-506-078-23

Align seq 1/1 to: US-09-506-078-23

178 CAGCATTGAGCTACGGCCTGCGCCCTGG

```
19 G|H|I|J|K|L|N|O|P|Q|R|S|T|U|V|W|X|Y|Z|
```

228 CTACGGCCTGCGGTCCGGGTGCTCTAGCC

```

32 rTvrG]vI@Ar@PrG]v
:
:
:

```

078

A	:
C	:
D	:
E	:
F	:
G	:
H	:
I	:
J	:
K	:
L	:
M	:
N	:
O	:
P	:
Q	:
R	:
S	:
T	:
U	:
V	:
W	:
X	:
Y	:
Z	:

```

328 TCTAGCTTCCCAAAAGTGGGCAAAAATATCTCTATT...CC 374
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
60 MetSerLeuPro...ThrProAlaProArgValThrValTyrValAspPr 75
375 CCAAAATTAC 384
|||
75 OProAlaTyr 78

```

seq_name: /cgn2_6/prodata/2/paa/US60_COMB.pep:US-60-120-454-23

```

seq_documentation_block:
; Sequence 23, Application US/60120454
; GENERAL INFORMATION:
; APPLICANT: Campos, Manuel
; APPLICANT: Yule, Terecila D
; APPLICANT: Martindod, Serge A
; APPLICANT: Durtsehl, Beeky A
; TITLE OF INVENTION: FUSION PROTEINS COMPRISING CARRIERS THAT CAN INDUCE A
; FILE REFERENCE: PC10202
; CURRENT APPLICATION NUMBER: US/60/120,454
; EARLIER APPLICATION NUMBER: N/A
; EARLIER FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 23
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 4GnRH-tmgd
; OTHER INFORMATION: encoded by pOE-GnRH.gd.
US-60-120-454-23

```

```

alignment_scores:
Quality: 208.50      Length: 70
Ratio: 4.088         Gaps: 5
Percent Similarity: 72.857   Percent Identity: 58.571

```

alignment_block:
US-09-306-689-12 x US-60-120-454-23 ..

Align seg 1/1 to: US-60-120-454-23 from: 1 to: 411

```

178 CAGCATTTGAGCTAGCGCCCTGCGCCGCTTCTCAAGATTGAG 227
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
19 GluHisTrpSerTyrGlyLeuArgProGly.....GluHisTrpSe 32
228 CTAGCGCGCTGCGGCTGCGCTAGCCAGATTGAGCTAGCGCGCTGC 277
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
32 TTYrGlyLeuArgProGly.....GluHisTrpSerTyrGlyLeuA 46
278 GCCCTGCGCGGCTAGCCAGATTGAGCTAGCGCGCTGCGCGGTGA 327
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
46 rProGly.....GluHisTrpSerTyrGlyLeuArgProGlySer 59
328 TCTAGCTTCCCAAAAGTGGGCAAAAATATCTCTATT...CC 374
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
60 MetSerLeuPro...ThrProAlaProArgValThrValTyrValAspPr 75
375 CCAAAATTAC 384
|||
75 OProAlaTyr 78

```

seq_name: /cgn2_6/prodata/2/paa/US086_COMB.pep:US-08-669-785-4

```

seq_documentation_block:
; Sequence 4, Application US/08669785
; GENERAL INFORMATION:
; APPLICANT: Betsou, Fotini
; APPLICANT: Sebou, Peter

```

```

; APPLICANT: Gulso, Nicole
; TITLE OF INVENTION: Protective Epitopes Of Adeny1
; TITLE OF INVENTION: Cyclase-Hemolysin(AC-Hly). Their Application To
; TITLE OF INVENTION: The Treatment Or To The Prevention Of Bordetella Infections
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/669,785
; FILING DATE: 27-JUN-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356, 0072-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1705 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-669-785-4

```

```

alignment_scores:
Quality: 206.50      Length: 372
Ratio: 1.059         Gaps: 11
Percent Similarity: 52.419   Percent Identity: 23.387

```

alignment_block:
US-09-306-689-12 x US-08-669-785-4 ..

Align seg 1/1 to: US-08-669-785-4 from: 1 to: 1705

```

300 TTGAGAGCTAGCGCCTGCGCTGCGGTGATCTAGCTCCAAAACGTGGG 349
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
380 LeuGluThrValProAlaSer.ProGlyLeuArgArgProSerLeuGlyA 396
350 CAAAAAATATCTCTATTCTCCCAAAATTCACATATGATACGAA 399
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
396 laValGlu.....ArgGlnAspSerGlyTyrAspSerLeu 407
400 CAAGTAAATGTTTACAGATTAGTCAAAAGCGCGCAAGAGTGGGAT 449
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
408 AspGlyValGlySerArgSerPhe..... 415
450 TGAGGTCAAAAGAGAAAGCAATTAATTGCAACAGCTCAACAGATT 499
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
416 .....SerL 417
500 TAGGCAGATTCAAAACCGCTATTGGCTTAAGTAGCGTGCGATTGTGTA 549
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
417 euGlyGluValSerAspMet.....AlaAlaVal 426
550 TCCGCTCCCAAAATGATTAATGCTACAGAAAACCTAAAGCGCCAGC 599
||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
427 GluAlaAlaGluLeuGluMetThrArgGlnValLeuHisAlaGlyAlaAr 443
600 ATTAGGTTCTGCCGAAAGCATTGTACAAAATGCAAAATGAAA.....GCCA 643

```

```

443  ggluAspAspAspGluProGlyValSerGlyAlaSerAlaHisTrpGlyG 693
444  AAACGTATTATTCGCAATTCATCTATTTTGGCTCAGTATTTGGCTGGA 693
460  lnrArgAlaLeuGlnGlyAlaGlnAlaValAlaAlaAlaAlaAlaAla 472
694  ATGGATTATTAGATGAGGCTTACAGAAATACAGCAACCAACATGCTTCG 743
473  ..... :GlnArgLeuValHisAlaIleAl 480
744  T..... AAAGTCGGCTTGGAGCTTACCAATTCATTA 775
480  aleuMetThrGlnPheGlyArgAlaGlySerThrAsnThrProGlnGlu 497
776  TTGAAAATATTCGTAATTCAGTAAAAACACTTGACGAAATTTGGAGCAA 825
497  lalAserLeuSerAlaAlaValAlaPheGlyLeuGlyGluAlaSerSerAl 513
826  ATTAGTCAA.....TTTGGTCCAAACTACAAATATCA 860
514  ValAlaGlnThrValSerGlyPhePheArgGlySerSerArgTrpAlaG 530
861  AGGCTTAGGAGCTTTAGGAGACAACCAACCAAAATATGGTGGAGCTGTA 910
530  yglPheGlyValAlaGlyGlyAlaMetAlaLeuGlyGlyGlyIleGly 547
911  AAGCTGGCCTTGGATTAGATTATCTCAAGGCTATTATCGGCGCAACA 960
547  lalValGlnAlaGlyMetSer..... 553
961  GCTGCACCTTGTACTTGCAGATAAAAATGCTTCAACAGCTAAAAAGTGG 1010
554  .....LeuThrAsp...AspAlaProIaGlyGlnTyAlaAl 565
1011  TGGGGGTTTGAATGGCAACAACGATTGTTGGTAATATTCACAAACGG 1060
565  aaIaGlyAlaGlnIleAlaLeuGlnLeuThrGlyGlyThrValGluLeu 582
1061  TTTCTTCTTACATTTTA.....GCCCAACGTTGGACGAGT 1098
582  lAserSerIleAlaLeuAlaLeuAlaAlaAlaAlaArgGlyValThrSerGly 598
1099  TTATCTTCAACGTGGCGCTGTGGCTGCTTAAATGCTTCTACTGTTTCT 1148
599  LeuGlnValAlaGlyAlaSerAlaGlyAlaAlaAlaGlyAlaLeuAlaAl 615
1149  TGGGATTAGCCCATTTAGCATTTGCCGGTATTCGCCGATAAATTTATCATG 1198
615  aaIaIaLeuSerProMetClnIleTyrGlyLeuValGlnGlnSerHisTyr 632
1199  CAAAAGTTTAGAGAGTTATGCGCAACGCTTAAAAAATTTAGGCTATGAC 1248
632  lAspGlnLeuAspIlysLeuAlaGlnGlnIleuSerSerAlaTyrGlyTyGlu 648
1249  GGAGATTAATTTATTCAGCAATATACAGGGGGAACAGGACTATTGATGC 1298
649  GlyAlaPheAlaLeuAlaGlnLeuTyrArgAspIysThrAlaIaGlnG 665
1299  ATCGGTACTGCAATTAATACCGCATTTGGCGCGCATATGCTGAGTGTGT 1348
665  yAlaValAlaGlyValSerAlaValLeuSerThrValGlyAlaAlaValS 682
1349  CTGCTGCTGCAGCG 1362
682  erIleAlaAlaAlaAla 686

```

```

1 TITLE INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
2
3 FILE REFERENCE: IGM-2004A0
4
5 CURRENT APPLICATION NUMBER: PCT/US00/22121
6
7 CURRENT FILING DATE: 2000-08-14
8
9 PRIOR APPLICATION NUMBER: 09/026,276
10
11 PRIOR FILING DATE: 1998-02-19
12
13 PRIOR APPLICATION NUMBER: US 09/374,721
14
15 PRIOR FILING DATE: 1999-08-13
16
17 NUMBER OF SEQ ID NOS: 35
18
19 SOFTWARE: PatentIn Ver. 2.0
20
21 SEQ ID NO 35
22
23 LENGTH: 40
24
25 TYPE: PRT
26
27 ORGANISM: Artificial Sequence
28
29 FEATURE:
30
31 OTHER INFORMATION: Description of Artificial Sequence: polypeptide
32
33 OTHER INFORMATION: antigen
34
35 PCT-US00-22121-35

```

```

alignment_scores:          Quality: 202.50          Length: 61
                           Ratio: 5.192           Gaps: 3
Percent Similarity: 63.934   Percent Identity: 63.934

alignment_block:
US-09-306-689-12 x PCT-US00-22121-35 ..
Align seg 1/1 to: PCT-US00-22121-35 from: 1 to: 40

1  CAGCATTTGAGCTACGAGCGGCTGCGGCTGCGAGCGGTTCTCAAGATTGGAG 74
   |||||||
1  GlnHisTrpSerTyrGlyLeuArgProGly..... 10

75  CTACGCGCTGCGGCTCGGGTGGCTGTACGACCAATTTGAGCTACGCGCTGC 122
   |||||||
1  .....GlnHisTrpSerTyrTlYLeuA 18

125  GCGCTGCGAGCGGTACGCCAAGATTGGAGCTACGCGCTGCGCGGATGGGA 174
   |||||||
18  rgrProGly.....GlnHisTrpSerTyrGlyLeuArgProGly... 30

175  TCTCAGCATTTGAGCTACGAGCGGCTGCGGCTGGC 207
   |||||||
31  ...GlnHisTrpSerTyrGlyLeuArgProGly 40

seq_name: /cgr12_6/prcodata/2/paa/US090_COMB.pep:US-09-026-276-35

seq_documentation_block:
1 Sequence 35, Application US/09026276
1 GENERAL INFORMATION:
1 APPLICANT: Kenten, John H
1 APPLICANT: Tramonitano, Alfonso
1 APPLICANT: Pilon, Aprile L
1 APPLICANT: Lohnas, Gerald L
1 APPLICANT: Roberts, Steven F
1 TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
1 FILE REFERENCE: U.S. Patent Application No. 09\026,276
1 CURRENT APPLICATION NUMBER: US/09/026,276
1 CURRENT FILING DATE: 1998-02-19
1 NUMBER OF SEQ ID NOS: 35
1 SOFTWARE: PatentIn Ver. 2.0
1 SEQ ID NO 35
1 LENGTH: 40
1 TYPE: PRT
1 ORGANISM: Porcine
1 US-09-026-276-35

alignment_scores:          Quality: 202.50          Length: 61
                           Ratio: 5.192           Gaps: 3
Percent Similarity: 63.934   Percent Identity: 63.934

```

alignment_block:

US-09-306-689-12 x US-09-026-276-35 ..

Align seg 1/1 to: US-09-026-276-35 from: 1 to: 40

```

25 CAGCATTTGAGACTAGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 74
   ||||||||||||||||||||||||||||||||||||||||||||
1  GlnHistrpserTyrglyLeuAArgProGly..... 10

75 CTAGCGCCCTGCGGTCCGGGTGCTCTAGCCAGCATTTGAGACTAGCGGCTGC 124
   ||||||||||||||||||||||||||||||||||||||||||||
11 .....GlnHistrpserTyrglyLeuA 18

125 GCCCTGGAGCGGTAGCCAGATTGAGCTAGGCGCTGCGGTGCGGTGGA 174
   |||||||| || ||||||||||||||||||||||||||||
18 rGProGly.....GlnHistrpserTyrglyLeuAArgProGly... 30

175 TCTCAGCATTTGAGCTAGGCGCTGCGCCCTGGC 207
   ||||||||||||||||||||||||||||||||||||||||||||
31 ...GlnHistrpserTyrglyLeuAArgProGly 40

```

seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:US-09-374-721A-35

seq_documentation_block:

```

; Sequence 35, Application US/09374721A
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H.
; APPLICANT: Roberts, Steven
; APPLICANT: Lohas, Gerald
; TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: CIP OF ICN-9601
; CURRENT APPLICATION NUMBER: US/09/374,721A
; CURRENT FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: 09/026,276
; PRIOR FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 35
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: polypeptide
; OTHER INFORMATION: antigen
US-09-374-721A-35

```

alignment_scores:

Quality:	202.50	Length:	61
Ratio:	5.192	Gaps:	3
Percent Similarity:	63.934	Percent Identity:	63.934

alignment_block:

US-09-306-689-12 x US-09-374-721A-35 ..

Align seg 1/1 to: US-09-374-721A-35 from: 1 to: 40

```

25 CAGCATTTGAGACTAGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 74
   ||||||||||||||||||||||||||||||||||||||||||||
1  GlnHistrpserTyrglyLeuAArgProGly..... 10

75 CTAGCGCCCTGCGGTCCGGGTGCTCTAGCCAGCATTTGAGACTAGCGGCTGC 124
   ||||||||||||||||||||||||||||||||||||||||||||
11 .....GlnHistrpserTyrglyLeuA 18

125 GCCCTGGAGCGGTAGCCAGATTGAGCTAGGCGCTGCGGTGCGGTGGA 174
   |||||||| || ||||||||||||||||||||||||||||
18 rGProGly.....GlnHistrpserTyrglyLeuAArgProGly... 30

175 TCTCAGCATTTGAGCTAGGCGCTGCGCCCTGGC 207
   ||||||||||||||||||||||||||||||||||||||||||||
31 ...GlnHistrpserTyrglyLeuAArgProGly 40

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:38:35 ; Search time 66.16 Seconds
(without alignments)
7.236 Million cell updates/sec

Title: US-09-306-689-2

Sequence: 58
1 XHMSYLRPGXXX 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_36:*

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	89.7	10	P20403	LH-RH analogue 1.
2	52	89.7	10	P40404	Sequence of peptide
3	52	89.7	10	P71316	Sequence of lutein
4	52	89.7	10	P71226	Luteinizing hormone
5	52	89.7	10	P71229	Luteinizing hormone
6	52	89.7	10	P71363	Sequence of a lute
7	52	89.7	10	P83021	Example of gonadol
8	52	89.7	10	R10699	LHRH analogue 3.
9	52	89.7	10	R15714	Peptide #2 with ho
10	52	89.7	10	R15715	Peptide #3 with ho
11	52	89.7	10	R29612	Goserelin, contg. a
12	52	89.7	10	R29703	Goserelin contg. a

13	52	89.7	10	18	W22391	Gonadotropin relea
14	52	89.7	10	18	W16391	Gonadotropin relea
15	51	87.9	9	20	W94891	LHRH peptide fragm
16	51	87.9	9	21	Y89609	Core polypeptide f
17	51	87.9	10	2	P10097	Sequence of lutein
18	51	87.9	10	2	P10411	Luteinizing Hormon
19	51	87.9	10	2	P10412	Luteinizing Hormon
20	51	87.9	10	2	P10416	Luteinizing Hormon
21	51	87.9	10	4	P30390	Gonadotropin-relea
22	51	87.9	10	4	P30391	Gonadotropin-relea
23	51	87.9	10	5	P40624	Gonadotropin-relea
24	51	87.9	10	5	P40629	Gonadolibertin anal
25	51	87.9	10	6	P50566	Sequence of gonado
26	51	87.9	10	6	P50222	Gonadotropin rele
27	51	87.9	10	7	P60127	Gonadolibertin anta
28	51	87.9	10	7	P61439	Sequence of lutein
29	51	87.9	10	7	P61403	Gonadotropin relea
30	51	87.9	10	7	P60576	Novel decapeptide
31	51	87.9	10	8	P70922	Luteinizing hormone
32	51	87.9	10	8	P70923	Luteinizing hormone
33	51	87.9	10	8	P71168	Luteinizing hormone
34	51	87.9	10	8	P71230	Luteinizing hormone
35	51	87.9	10	8	P71477	Luteinizing hormone
36	51	87.9	10	8	P71492	LHRH peptide anal
37	51	87.9	10	10	P90630	Sequence of lutein
38	51	87.9	10	11	R06475	Gonadotropin relea
39	51	87.9	10	12	R10507	Pamatoe of lutein
40	51	87.9	10	12	R14140	LHRH analogue #1 c
41	51	87.9	10	12	R14141	LHRH analogue #2 c
42	51	87.9	10	12	R14672	Generic peptide w/
43	51	87.9	10	12	R15713	Peptide #1 with ho
44	51	87.9	10	12	R15717	Peptide #5 with ho
45	51	87.9	10	21	Y55061	Luteinizing hormone

ALIGNMENTS

RESULT 1	
P20403	
ID	P20403 standard; Protein: 10 AA.
AC	P20403;
XX	
DT	01-DEC-1992 (first entry)
XX	
DE	LH-RH analogue 1.
XX	
KW	Luteinizing hormone/releasing hormone; microcapsule; hydrolysis:
XX	fertility.
OS	Synthetic.
XX	
FT	Key
FT	Modified-site 1
FT	Location/Qualifiers
FT	/label= OTHER
FT	/note= "Pyroglutamic acid"
FT	6
FT	/label= OTHER
FT	/note= "3-(2-naphthyl)-D-Ala"
XX	
PN	EP52510-A.
XX	
PD	26-MAY-1982.
XX	
PF	17-NOV-1981; 81EP-0305426.
XX	
PR	18-NOV-1980; 80US-0207864.
PR	11-DEC-1981; 81US-0329832.
XX	
PA	(SYNT) SYNTX USA INC.
XX	
PI	Kent JS, Sanders LM, Lewis DH, Tice TR;

XX WPI: 1982-44213E/22 (44213E).
 DR Microencapsulated polypeptide sustained release compsn. - useful
 XX for release of luteinizing hormones etc. for fertility control
 PT
 XX Disclosure: Page 7: 27pp; English.
 PS
 XX The sequence in P20121 is the consensus sequence of a hormonally
 CC active polypeptide. The sequences in P20403-6 are specific examples of
 CC these polypeptides. These polypeptides are water soluble and are
 CC analogues of luteinizing hormone/releasing hormone (LH-RH). These
 CC peptides are used within the scope of the invention in hormonally
 CC active microcapsule formulations comprising at least one of these
 CC polypeptides and optionally a polymer hydrolysis modifying agent
 CC intimately mixed with or coated by a biocompatible, biodegradable
 CC polymer which when administered to a mammal, will release a daily
 CC amount of polypeptide effective for maintaining an hormonally
 CC related condition over a predetermined period of time. The
 CC specific hormonally related condition is the control of fertility
 CC and physiologically related effects.
 CC
 XX Sequence 10 AA:
 SQ

Query Match 89.7%; Score 52; DB 3; Length 10;
 Best Local Similarity 88.9%; Pred. No. 0.0011;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
 |||||
 DB 2 hwsyalrpg 10

RESULT 2
 P40404 P40404 standard; peptide; 10 AA.
 AC P40404;
 XX
 XX 04-FEB-1992 (first entry)
 DT
 XX Sequence of peptide in a nasal composition having luteinizing
 DE hormone releasing hormone (LHRH) agonist activity.
 XX

KW Abortion; inducer; fertility control; growth promoter;
 chemical sterilant; prostate cancer therapy; hormone.
 XX

Key Location/Qualifiers
 FT Modified-site 1
 FT Modified-site 6 /label= pyroglu
 FT Modified-site 10 /label= 3-(2-naphthyl)-D-Ala
 FT Modified-site 10 /label= Gly-NH2
 FT

EP111841-A.
 PN 27-JUN-1984.
 PD 08-DEC-1983; 83EP-0112369.
 PF 10-DEC-1982; 82US-0448548.
 PR (SYNT) SYNTX (USA) INC.
 XX
 PA Anik ST;
 XX
 PI WPI: 1984-159888/26.
 DR
 XX
 PT Nasal compsn. contg. nona: or deca:peptide - having LHRH
 PT activity, with bile acid (salt) surfactant for enhanced
 PT absorption
 XX

XX Claim 4; Page 24; 33pp; English.
 PS
 XX The inventors claim a nasal compsn. having LHRH agonist activity,
 CC together with a bile acid or its salt as surfactant, in aq. soln.
 CC prefd. peptides are described in US4234571 and are useful for
 CC fertility control, growth enhancement, treatment of prostatic cancer,
 CC inducing abortion. The compsn. comprises 0.005-5 mg/ml of the
 CC peptide, with 0.2-5% w/v of the surfactant, together with the buffer
 CC soln. A typical compsn. contains 0.05-4 mg/ml of a peptide and 0.5-4%
 CC sodium glycocholate. The molar ratio of bile acid to peptide is
 CC equal to or greater than 20:1. An additional surfactant may be
 CC included.
 XX
 SO Sequence 10 AA:

Query Match 89.7%; Score 52; DB 5; Length 10;
 Best Local Similarity 88.9%; Pred. No. 0.0011;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
 |||||
 DB 2 hwsyalrpg 10

RESULT 3
 P71316 P71316 standard; peptide; 10 AA.
 AC P71316;
 XX

DT 30-APR-1991 (first entry)
 XX

DE Sequence of luteinizing hormone releasing hormone (LHRH) analogues,
 DE having D-3 pyridyl-alanine at position 6.
 XX

KW Prostate cancer therapy; agonist.
 XX
 OS Homo sapiens.
 XX

Key Location/Qualifiers
 FT Misc-difference 1 /label= pyro-glu
 FT Misc-difference 6 /label= 3-(3-pyridyl)-D-Ala
 FT Misc-difference 9 /label= Pro-NHCH3, Pro-NHC2H5, Pro-NHC3H7
 FT Misc-difference 10 /label= "If Gly (10) absent"
 FT Misc-difference 10 /label= Gly-NH2
 FT /note= "or absent"
 FT

US4642332-A.
 PN 10-FEB-1987.
 PD 26-APR-1985; 85US-0727711.
 PF 26-APR-1985; 85US-0727711.
 PR (TEXA) UNTV TEXAS SYSTEM.
 XX
 PA Folkers K, Jie-Cheng X, Bowers CY;
 XX
 PI WPI: 1987-056353/08.
 DR
 XX
 PT Luteinizing hormone releasing hormone analogues - of greater
 PT potency, having D-3-pyridyl-alanine or 3-quinolyl-D-alanine at
 PT position 6
 PS Claim 1; column 11; 7pp; English.
 XX

CC These peptides are synthetic analogues of LHRH which are termed
 CC super agonists because of their superior potency. They may be used
 CC to treat prostate cancer. The peptides are synthesised by solid
 CC phase techniques using benzhydrylamine resin as solid support.

XX Sequence 10 AA;

Query Match 89.7%; Score 52; DB 8; Length 10;

Best Local Similarity 88.9%; Pred. No. 0.0011; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
 |||||
 Db 2 hwsyalrpg 10

RESULT 4

P71226 P71226 standard; peptide; 10 AA.

XX AC P71226;

XX DT 29-APR-1991 (first entry)

XX DE Luteinising hormone releasing hormone analogue.

XX KW Luteinising hormone releasing hormone; (D-Ala6)-LHRH; cancer;
 XX KM cytotoxic effects; bone marrow; agonist; radiotherapy; chemotherapy;

XX PN CH659392-A.

XX PD 30-JAN-1987.

XX PF 07-MAY-1984; 84CH-0002226.

XX PR 07-MAY-1984; 84CH-0002226.

XX PA (DEBI-) DEBIOPHARM SA.

XX PI Mathe G, Misset JL, Lemaigre G, Mauvernay RY, Schally AV;

XX DR WPI; 1987-037621/06.

XX PT Compsns. inhibiting cytotoxic effects in cancer therapy - contain
 PT luteinising and follicle stimulating hormone release hormone or
 PT its analogues

XX PS Claim 2; page 2; 4pp; English.

XX CC The analogue is an agonist. It inhibits the cytotoxic effects of
 CC cancer therapy on bone marrow cells, and protects the cells from the
 CC undesirable effects of radio- and chemotherapy used to treat hormone-
 CC dependent and -independent cancers. The compsns. may be in sustained
 CC release form and may be given parenterally, esp. i.m.

XX SQ Sequence 10 AA;

Query Match 89.7%; Score 52; DB 8; Length 10;
 Best Local Similarity 88.9%; Pred. No. 0.0011;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
 |||||
 Db 2 hwsyalrpg 10

RESULT 5

P71229 P71229 standard; protein; 10 AA.

XX AC P71229;

XX XX 29-APR-1991 (first entry)

XX DT Luteinising hormone releasing hormone analogue.

XX DE Luteinising hormone releasing hormone; (D-Ser(Bu+6,des Gly-NH2 10)-LHRH;

XX KW cytotoxic effects; bone marrow; agonist; radiotherapy;

XX KM chemotherapy; cancer.

XX PN CH659392-A.

XX PD 30-JAN-1987.

XX PF 07-MAY-1984; 84CH-0002226.

XX PR 07-MAY-1984; 84CH-0002226.

XX PA (DEBI-) DEBIOPHARM SA.

XX PI Mathe G, Misset JL, Lemaigre G, Mauvernay RY, Schally AV;

XX PT Compsns. inhibiting cytotoxic effects in cancer therapy - contain
 PT luteinising and follicle stimulating hormone release hormone or
 PT its analogues

XX PS Claim 2; page 2; 4pp; English.

XX CC The analogue is an agonist. It inhibits the cytotoxic effects of
 CC cancer therapy on bone marrow cells, and protects the cells from the
 CC undesirable effects of radio- and chemotherapy used to treat hormone-
 CC dependent and -independent cancers. The compsns. may be in sustained
 CC release form and may be given parenterally, esp. i.m.

XX SQ Sequence 10 AA;

Query Match 89.7%; Score 52; DB 8; Length 10;
 Best Local Similarity 88.9%; Pred. No. 0.0011;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
 |||||
 Db 2 hwsyalrpg 10

RESULT 6

P71363 P71363 standard; peptide; 10 AA.

XX AC P71363;

XX DT 26-MAY-1991 (first entry)

XX DE Sequence of a luteinising hormone-releasing hormone (LHRH) analogue
 DE used in a compsn. to treat infertility.

XX KW Infertility treatment; impotency; spermatogenesis; androgen;
 XX sterlant abortion.

XX FH Key Location/Qualifiers

FT MISC-difference 1 /Label= (PYRO)Glu

FT MISC-difference 5 /Label= Tyr-3-(naphthyl)

FT MISC-difference 6 /Label= D-Ala

FT MISC-difference 10 /Label= Gly-NH2

XX FT US4675189-A.

PD 23-JUN-1987.
 XX
 PF 08-FEB-1985; 85US-0699715.
 XX
 PR 11-DEC-1981; 81US-0329832.
 PR 18-NOV-1980; 80US-0207864.
 PR 08-FEB-1985; 85US-0699715.
 XX
 PA (SYNTE) SYNTEX (USA) INC.
 PI Kent JS, Lewis DH, Sanders LM, Tice TR;
 DR WPI; 1987-192071/27.
 XX
 PT Sustained release compsn. microcapsules - contain luteinising
 PT hormone releasing hormone analogues and
 PT (poly-lactide-co-glycolide) polymer, to treat infertility
 XX
 PS Claim 9; column 18; 11pp; English.
 XX
 CC Low doses of the compsn. are used to treat infertility in females
 CC and impotency and to stimulate spermatogenesis and androgen prodn.
 CC in males. High doses block ovulation and suppress spermatogenesis;
 CC and may be used to promote wt. gain in animals, to stimulate
 CC abortion in animals and as a chemical sterilant. (See US4234571).
 CC For agonist fertility control, the microcapsules release the drug to
 CC dose 0.01-100 (pref. 0.1-5.0) micrograms/kg/day.
 CC
 SQ Sequence 10 AA;

Query Match 89.7%; Score 52; DB 8; Length 10;
 Best Local Similarity 88.9%; Pred. No. 0.0011;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 HWSYXLRPG 10
 | | | | | | | |
 DB 2 hwsyalrpg 10

RESULT 7
 P83021
 ID P83021 standard; protein; 10 AA.
 XX
 AC P83021;
 XX
 DT 11-DEC-1990 (first entry)
 XX
 DE Example of gonadoliberin peptide analogue.
 XX
 KM Gonadoliberin; agonists; antagonists; fertility; amenorrhoea;
 KM oligospermia; contraceptives; cancers.
 XX
 OS synthetic.
 OS
 PN EP263521-A.
 XX
 PD 13-APR-1988.
 XX
 PF 08-OCT-1987; 87EP-0114702.
 XX
 PR 09-OCT-1986; 86DE-3634435.
 XX
 PA (FARH) HOECHST AG.
 PI Konig W, Sandow JK, Kolar C;
 DR WPI; 1988-099948/15.
 XX
 PT New peptide derivs. contg. glycosyl gp. - useful as gonadoliberin
 PT agonists or antagonists
 XX
 PS Claim 1; Page 40; 43pp; German.

XX This peptide is an agonist or antagonist of gonadoliberin. It, and
 CC its salts, is useful as an agonist for promoting fertility e.g. for
 CC treating amenorrhoea, luteal insufficiency, oligospermia and retarded
 CC puberty. It is also useful as a contraceptive and for treating
 CC steroid-dependent cancers. As an antagonist it is useful for
 CC treating gonadotropin- and steroid-dependent diseases.
 CC The N-terminal has an H gp. or a 1-7C acyl attached. Residues 1, 2
 CC and 3 are opt. ring subst. by one or two of Br, Cl, F, NO2, NH2, Me
 CC or OMe. Residues 6 and 7 opt. have a partially protected glycosyl
 CC attached which increases their water-solubility without impairing
 CC their activity.
 CC See also P82831 and P83017-20.
 CC
 SQ Sequence 10 AA;

Query Match 89.7%; Score 52; DB 9; Length 10;
 Best Local Similarity 88.9%; Pred. No. 0.0011;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 HWSYXLRPG 10
 | | | | | | | |
 DB 2 hwsyalrpg 10

RESULT 8
 R10699
 ID R10699 standard; peptide; 10 AA.
 XX
 AC R10699;
 XX
 DT 22-APR-1991 (first entry)
 XX
 DE LHRH analogue 3.
 XX
 KM Luteinising hormone releasing hormone.
 XX
 OS Synthetic.
 OS
 PN EP413209-A.
 XX
 PD 20-FEB-1991.
 XX
 PF 02-AUG-1990; 90EP-0114829.
 XX
 PR 10-JUL-1990; 90US-0548512.
 PR 07-AUG-1989; 89US-0390572.
 XX
 PA (ABBO) ABBOTT LABORATORIES.
 XX
 PI Oxford R, Greer J;
 DR WPI; 1991-052535/08.
 XX
 PT New pseudo nona- and deca-peptide derivs. of LHRH - used to
 PT modulate level of sex hormones in males and females
 XX
 PS Claim 4; Page 74; 78pp; English.

CC The peptide is one of several examples of compounds derived from the
 CC sequence of LHRH in which the nitrogen atom of at least one of the
 CC amide bonds has been alkylated. The peptides are used to modulate
 CC levels of sex hormones in males and females. They are stable
 CC against proteolytic enzymes and biologically potent after oral admin.

CC in animals and humans. See also R10697-R10698.
 XX Sequence 10 AA:
 SQ

Query Match 89.7%; Score 52; DB 12; Length 10;
 Best Local Similarity 88.9%; Pred. No. 0.0011;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
 |||||
 Db 2 hwsyalrpg 10

RESULT 9

R15714 ID R15714 standard; Protein; 10 AA.

AC R15714;

DT 24-JAN-1992 (first entry)

DE Peptide #2 with homology to LHRH.

XX Luliberin.

XX Synthetic.

OS Key Location/Qualifiers

FH Modified-site 1

FT /label= OTHER

FT /note= "pyroglu"

FT Modified-site 6

FT /label= OTHER

FT /note= "D-Ala"

FT Modified-site 9

FT /label= Hyp

FT Modified-site 10

FT /label= OTHER

FT /note= "amidated"

PN WO9116343-A.

PD 31-OCT-1991.

PF 22-APR-1991; 91WO-FR00332.

PR 23-APR-1990; 90FR-0005147.

PA (INRM) INSERM INST NAT SANTE.

DR Gautron J, Patton E, Kordon C, Bauer K;

WP1: 1991-339753/46.

PS New peptide homologous with luteinising hormone-releasing hormone

PT - used to treat gynaecological conditions, cancer of gonads and

XX sec. sexual organs, psychiatric conditions and in assays

CC Claim 6; Page 50; 83pp; French.

CC The C-terminal residue (Gly-CO-NH2) can be replaced by ethylamide.

CC This peptide and fragments of it (i.e. amino acids 4-10, 5-10, 6-10

Query Match 89.7%; Score 52; DB 12; Length 10;
 Best Local Similarity 88.9%; Pred. No. 0.0011;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
 |||||
 Db 2 hwsyalrpg 10

RESULT 10

R15715 ID R15715 standard; Protein; 10 AA.

AC R15715;

DT 24-JAN-1992 (first entry)

DE Peptide #3 with homology to LHRH.

XX Luliberin.

XX Synthetic.

OS Key Location/Qualifiers

FH Modified-site 1

FT /label= OTHER

FT /note= "pyroglu"

FT Modified-site 6

FT /label= OTHER

FT /note= "D-Ser"

FT Modified-site 9

FT /label= Hyp

FT Modified-site 10

FT /label= OTHER

FT /note= "amidated"

PN WO9116343-A.

PD 31-OCT-1991.

PF 22-APR-1991; 91WO-FR00332.

PR 23-APR-1990; 90FR-0005147.

PA (INRM) INSERM INST NAT SANTE.

DR Gautron J, Patton E, Kordon C, Bauer K;

WP1: 1991-339753/46.

PS New peptide homologous with luteinising hormone-releasing hormone

PT - used to treat gynaecological conditions, cancer of gonads and

XX sec. sexual organs, psychiatric conditions and in assays

CC Claim 6; Page 50; 83pp; French.

CC The C-terminal residue (Gly-CO-NH2) can be replaced by ethylamide.

CC This peptide and fragments of it (i.e. amino acids 4-10, 5-10, 6-10

CC and 7-10) may be agonists or antagonists of LHRH. They are useful for

CC treating e.g. precocious or delayed puberty, psychiatric disorders

CC esp. those of the libido or sexual aggression, etc. in addition they

CC are useful for functional exploration of the hypothalamus-hypophyseal

CC axis and for radioimmunological or biological assay (of LH, FSH and

CC steroid levels) in biological fluids and biopsy samples.

CC Sequence 10 AA:

Query Match 89.7%; Score 52; DB 12; Length 10;
 Best Local Similarity 88.9%; Pred. No. 0.0011;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10

Db 2 hwsyslrpg 10

RESULT 11

R29612 ID R29612 standard; peptide; 10 AA.

AC R29612;

DT 23-APR-1993 (first entry)

DE Goserelin, contg. C-terminal aza-amino acid amide.

KW LHRH; luteinising hormone releasing hormone; analogue; prostate; cancer; breast; gynaecological conditions.

OS Synthetic.

FT Key Location/Qualifiers

FT Modified-site 1 /note= "pyroglu"

FT Misc-difference 6 /note= "D-Ser(But)"

FT Modified-site 10 /note= "amidated azagly"

EP518656-A.

PD 16-DEC-1992.

PF 11-JUN-1992; 92EP-0305341.

PR 14-JUN-1991; 91GB-0012859.

PA (ICIL) IMPERIAL CHEM IND PLC.

XX Hayward CF;

DR WPI; 1992-417511/51.

PT Solid phase synthesis of peptide(s) contg. aza-aminoacid - partic. goserelin, by assembling aminoacid(s) of peptide except C terminal aza aminoacid and cleaving peptide from support with hydrazine, etc.

PS Disclosure; Page 2; 4pp; English.

XX Goserelin contg. a C-terminal aza-amino acid amide was prepd. by a novel method comprising assembling all the amino acids of the peptide except the C-terminal aza-amino acid by conventional solid phase synthesis, cleaving the peptide from the support with hydrazine and reacting the released hydrazine with a cyanate ion. The peptide also contains a tert-butyloxy gp. in its side chain. CC The peptide is an LHRH analogue useful in the treatment of prostate cancer, breast cancer and certain gynaecological conditions.

XX Sequence 10 AA;

Query Match 89.7%; Score 52; DB 13; Length 10;

Best Local Similarity 88.9%; Pred. No. 0.0011; Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10

Db 2 hwsyslrpg 10

RESULT 12

R29203

ID R29703 standard; peptide; 10 AA.

AC R29703;

DT 23-APR-1993 (first entry)

DE Goserelin contg. aza-amino acid amide.

KW Ester; LHRH; luteinising hormone releasing hormone; prostate; cancer; breast; gynaecological conditions.

OS Synthetic.

FT Key Location/Qualifiers

FT Modified-site 1 /label= pglu

FT Misc-difference 6 /note= "D-Ser(But)"

FT Modified-site 10 /note= "amidated azagly"

EP518655-A.

PD 16-DEC-1992.

PF 11-JUN-1992; 92EP-0305340.

PR 14-JUN-1991; 91GB-0012825.

PA (ICIL) IMPERIAL CHEM IND PLC.

XX Cotton R, Giles MB;

DR WPI; 1992-417510/51.

PT New N-protected-aza-amino-acyl ester(s) - are used in solid phase synthesis of peptide(s) contg. C-terminal aza-aminoacid, partic. goserelin

PS Disclosure; Page 2; 7pp; English.

XX Goserelin contg. an aza amino acid was prepd. by a novel solid phase synthesis method comprising reacting an active ester or imidazolone of an N-protected aza-amino acid with an appropriate reactive solid support, carrying out further conventional solid phase peptide synthesis steps to add sequentially further amino acids, to form a peptide with the required amino acid sequence bound to a solid support, cleaving the peptide from its support and opt. reacting the prod. with hydrazine to remove any acyl gps. formed during synthesis. CC The peptide is a synthetic analogue of LHRH and is useful in the treatment of prostate cancer, breast cancer and certain gynaecological conditions.

XX Sequence 10 AA;

Query Match 89.7%; Score 52; DB 13; Length 10;

Best Local Similarity 88.9%; Pred. No. 0.0011; Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10

Db 2 hwsyslrpg 10

RESULT 13

W22391

ID W22391 standard; peptide; 10 AA.

AC W22391;

DT 20-JAN-1998 (first entry)

DE Gonadotropin releasing hormone [(3-Mercaptopropanoyl)-Gln1,dala6]GnRH.

KW	Gonadotropin releasing hormone; vaccine; immuno-sterilisation;
RW	fertility control; steroid hormone stimulated tumour; endometriosis;
KM	GnRH.
XX	
OS	Synthetic.
XX	
FH	Key
FT	Modified-site
FT	Misc-difference
FT	Modified-site
XX	
PN	WO9715316-A1.
XX	
PD	01-MAY-1997.
XX	
PE	23-OCT-1996; 96WO-US16950.
XX	
PR	13-FEB-1996; 96GB-0002951.
PR	27-OCT-1995; 95US-0005905.
XX	
PA	(MERI) MERCK & CO INC.
XX	
PI	Lombardo VK, Marburg S, Tolman RL;
XX	
DR	WPt; 1997-258759/23.
PT	New gonadotropin releasing hormone conjugates - used as vaccines for
PT	immuno-sterilisation of animals, fertility control or for treating
PT	steroid hormone stimulated tumours or conditiions.
XX	
PS	Disclosure; Page 33; 75pp; English.
XX	
CC	A gonadotropin releasing hormone conjugate has been developed which has
CC	the formula: Z-[L1-Am-(Y1(L2-X)-Ar)q-Y2(L2-X)-Ap.OH]r, where A = an
CC	amino acid selected from Gly, Ser, Thr, beta-Ala, and Ala, with the
CC	proviso that at least one A is Ser or Thr; L1 = a linker optionally
CC	attached to an internal marker; L2 = a linker; X = a gonadotrophin
CC	releasing hormone (GnRH) having a free sulphhydryl group; Y1, Y2 = Lys
CC	or ornithine (Orn); Z = an immunogenic carrier protein; m = 0-3; n = 5
CC	or 6; p = 0-1; q = 1 or 2; r = 1-3. They are useful as vaccines
CC	for use in immunosterilising animals, fertility control, and for
CC	treatment of steroid hormone stimulated tumour or conditions such as
CC	endometriosis. The conjugates employ a hydrophilic linear scaffold as a
CC	means to increase loading of desirable hapten onto a carrier protein.
CC	Compared to conventional GnRH conjugates, these provide more GnRH per
CC	carrier molecule while utilising a minimal number of carrier molecule
CC	attachment sites. The present sequence represents [(3-Mercaptopropionyl)
CC	-Gln1,dAla6]GnRH which is used in an example of the present invention.
CC	The novel vaccine compositions provide a high titre of anti-GnRH
CC	antibody and produce minimal injection site lesions.
XX	
SO	Sequence 10 AA:
OY	Query Match 89.7%; Score 52; DB 18; Length 10; Best Local Similarity 88.9%; Pred. No. 0.0011; Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
DB	2 HWSXYLRPG 10 2 hwsyalrpg 10
RESULT 14	
ID W16391	W16391 standard; peptide; 10 AA.
NC W16391;	
DT 14-JAN-1998	(first entry)

DE Gonadotropin releasing hormone analogue, preparation 4.

XX
XX
KW Gonadotropin releasing hormone; immunogen; Pseudomonas exotoxin;
antibody; fertility; aggressive behaviour; tumour; steroid hormone;
KW endometriosis; vaccine; sterilise; breast cancer; prostatic cancer.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FH Modified-site 1
FT /note= "Attached to the linker HSCGH2CH2CONH--"
FT Misc-difference 6 /note= "D-form residue"
FT Modified-site 10
FT /note= "In amide form"
XX
XX
PN WO9715325-A1.
XX
PD 01-MAY-1997.
XX
XX 23-OCT-1996; 96WO-US17008.
XX
XX 13-FEB-1996; 96GB-0002878.
PR 27-OCT-1995; 95US-0008018.
XX
XX (MERI) MERCK & CO INC.
XX
XX Hickey GJ, Mohn KL;
PI
XX WPI; 1997-258766/23.
DR
XX
XX Immunogen comprising gonadotropin releasing hormone and Pseudomonas
PT exotoxin - generates antibodies that control fertility, reduce
PT aggressive behaviour, and treat tumours responsive to steroid
XX hormones or endometriosis
XX
XX
XX Disclosure; Page 34; 63pp; English.
PS
XX
XX A method has been developed for the generation of anti-gonadotropin
CC releasing hormone (GnRH) antibodies (Ab) by administration of an
CC immunogen consisting of GnRH associated with Pseudomonas exotoxin (PE)
CC or its variants. The present sequence represents a GnRH analogue,
CC preparation 4, used in the preparation of GnRH vaccines
CC containing this immunogen are used to control fertility; reduce
CC undesirable behaviour (e.g. aggression) induced by reproductive
CC hormones and to treat (a) tumours that respond to steroid sex hormones
CC (e.g. breast and prostatic cancers) or (b) endometriosis. Particularly
CC the immunogen is used to sterilise animals, e.g. to improve feed
CC efficiency and carcass quality, and to eliminate boar taint. PG is an
CC effective immunogenic carrier for GnRH, providing high titres of Ab.
XX
XX Sequence 10 AA:
SO

Query Match 89.7%; Score 52; DB 18; Length 10;
Best Local Similarity 88.9%; Pred.No. 0.0011;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0.

QY 2 HWSYXLRP6 10
| | | | |
DB 2 HWSYALRP9 10

RESULT 15
W94891
1D. W94891 standard; peptide; 9 AA.
XX
AC W94891;
XX
DT 11-MAY-1999 (first entry)
XX
XX LHRH peptide fragment.

```

XX LHRH; immune response; luteinising hormone releasing hormone; DT;
KM diphtheria toxoid; castrating; oestrus cycling; aggression; breast;
KM sexual activity; organoleptic; livestock; cell growth; malignant;
KM prostate; ovarian; oncofoetal; hyperplastic; pregnancy;
KM endometriosis; inflammatory response.
XX
OS Homo sapiens.
XX
PN W09902180-A1.
XX
PD 21-JAN-1999.
XX
PF 09-JUL-1998; 98W0-AU00532.
XX
PR 09-JUL-1997; 97AU-0007768.
XX
PA (CSLC-) CSL LTD.
XX
PI McNamara MK;
XX
DR WPI; 1999-120511/10.
XX
PT New immunogenic leutenising hormone releasing hormone compositions -
PT comprise LHRH conjugated to diphtheria toxoid and adsorbed to an
PT ionic polysaccharide, used to inhibit reproductive function in
PT animals
XX
PS Example 3; Page 30; 41pp; English.
XX
CC The invention relates immunogenic composition for eliciting an immune
CC response to luteinising hormone releasing hormone (LHRH). The composition
CC comprises a LHRH-diphtheria toxoid (DT) conjugate adsorbed to an ionic
CC polysaccharide. The LHRH-DT compositions can be used for eliciting an
CC immune response to LHRH, for castrating an animal, for regulating an
CC oestrus cycling in a female animal or for inhibiting characteristics
CC induced by the sexual maturation of an animal, e.g. aggression or sexual
CC activity. They can also be used for achieving production gains in
CC livestock, e.g. reduction or elimination of unwanted organoleptic
CC characteristics from the meat of livestock. They can also be used for
CC inhibiting the growth of cells which are regulated directly or indirectly
CC by LHRH, e.g. malignant breast cells, malignant prostate cells, malignant
CC ovarian cells, malignant oncofoetal cells or hyperplastic cells. They can
CC also be used for down-regulating the libido of an animal. They can also
CC be used for inhibiting pregnancy, prostate enlargement, endometriosis or
CC inflammatory responses. The LHRH compositions induce a more effective
CC immune response against LHRH than the LHRH-carrier adjuvant compositions.
CC The effective immune response against LHRH results in prevention of the
CC release of the hormones LH and FSH from the anterior pituitary. Sequences
CC W94890-93 are peptide derivatives of LHRH.
XX
SO Sequence 9 AA:

```

```

Query Match      87.9%; Score 51; DB 20; length 9;
Best local similarity 88.9%; Pred. No. 2.1e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 HWSYXLRPG 10
   ||||| ||||
Db 1 hwsyglrpg 9

```

Search completed: March 2, 2001, 10:53:45
Job time: 910 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:53:47 ; Search time 66.16 Seconds
(without alignments)
359.200 Million cell updates/sec

Title: US-09-306-689-13

Perfect score: 3591
Sequence: 1 MATVIDRSGHWSYGLRPGSG.....GLRPGSGSDMSYGLRPGGS 695

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_36: *
1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT: *
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT: *
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT: *
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT: *
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT: *
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT: *
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT: *
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT: *
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT: *
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT: *
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT: *
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT: *
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT: *
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT: *
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT: *
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT: *
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT: *
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT: *
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT: *
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT: *
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3591	100.0	695	19	LKT-GNRH chimeric
2	3591	100.0	695	21	Leukotoxin/gonadot
3	3591	100.0	695	21	Gonadotropin relea
4	2642.5	73.3	544	19	LKT-GNRH chimeric
5	2633.5	73.3	544	17	LKT-GNRH chimeric
6	2426	67.6	977	17	LKT-GNRH chimeric
7	2426	67.6	977	19	LKT-GNRH chimeric
8	2354	65.6	490	20	Leukotoxin carrier
9	2173	60.5	936	14	GNRH-Leukotoxin ge
10	2131	59.3	1069	15	Bovine IFNgamma/Lk
11	2131	59.3	1069	18	Chimeric protein #
12	2126	59.2	926	12	LKT352: Pasteurel

13	2126	59.2	926	14	R34545
14	2126	59.2	926	15	R50291
15	2126	59.2	926	17	W03945
16	2126	59.2	926	19	W79568
17	2117.5	59.0	1098	13	R22103
18	2117.5	59.0	1098	15	R52747
19	2117.5	59.0	1098	18	W13866
20	2116	58.9	924	14	R42385
21	2116	58.9	924	14	R42380
22	2116	58.9	924	14	R42378
23	2112	58.8	943	14	R34546
24	2110	58.8	951	14	R34548
25	2108.5	58.7	953	11	R07167
26	2108.5	58.7	953	14	R43885
27	2108.5	58.7	953	15	R60072
28	2105.5	58.6	953	12	R15159
29	2103	58.6	924	12	R10889
30	1650.5	46.0	1334	15	R50290
31	1650.5	46.0	1403	19	W79574
32	1644.5	45.8	1334	12	R14481
33	1551.5	43.2	1403	12	R10890
34	1524	42.4	934	17	W07637
35	1264.5	35.2	956	12	R12561
36	1264.5	35.2	956	18	W22156
37	1264.5	35.2	956	21	W51410
38	938.5	26.1	1049	18	W22159
39	938.5	26.1	1049	21	W51412
40	938.5	26.1	1244	15	R54781
41	848.5	23.6	1022	18	W22152
42	848.5	23.6	1022	21	W51406
43	828.5	23.1	1023	16	R76991
44	403	11.2	758	17	R86998
45	380.5	10.6	323	12	R11187

ALIGNMENTS

RESULT 1	W79573	standard; Protein; 695 AA.
ID	W79573	
AC	W79573:	
DT	24-DEC-1998	(first entry)
XX	LKT-GNRH chimeric protein.	
DE	Chimeric: PCB122; LKT 111; GNRH: Gonadotropin releasing hormone; multimeric	
KW	cytotoxic activity; antigen presentation; immune response; vaccine;	
KW	tumour.	
XX	Synthetic.	
OS	WO9806848-A1.	
XX	19-FEB-1998.	
PD	08-AUG-1997:	97WO-CA00559.
PF	09-AUG-1996:	96US-0694865.
XX	(UYSA-) UNIV SASKATCHEWAN.	
PA	Manns JG, Potter AA;	
PI	WPI: 1998-159540/14.	
XX	N-PSDB: V61535	
DR	Chimeric protein of leukotoxin and gonadotropin releasing hormone	
XX	useful for, e.g. preparation of vaccines for reduction of incidence	
PT	of mammary tumours in mammals	
XX		

PS- Claim 9; Figure 9.1-6; 118bp; English.

CC The present sequence represents the LKT-GnRH chimeric protein from
 CC PC8122. This plasmid contains the LKT 111 polypeptide fused to sixteen
 CC copies of the GnRH peptide, in the pattern of: 8 copies of GnRH-LKT 111-8
 CC copies of GnRH. This chimera lacks cytotoxic activity which enables
 CC there to be an increase in antigen presentation and thus an optimal
 CC immune response. The removal of this region also enables the truncated
 CC LKT to be expressed at much higher levels and allows the amount of
 CC antigen administered to be reduced. This chimeric protein comprises a
 CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The
 CC chimeric protein can be used as a vaccine to help reduce the incidence of
 CC mammary tumours in a mammalian individual.

XX Sequence 695 AA:

Query Match 100.0%; Score 3591; DB 19; Length 695;
 Best Local Similarity 100.0%; Pred. No. 2.7e-233;
 Matches 695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATVIDRSOHMSYGLRPGSGSDMSYGLRPGSSOHMSYGLRPGSGSDMSYGLRPGSGS
 DB 1 macvdrsqmwsyglirpssgsqdwsgylirpssgsqdwsgylirpssgsqdwsgylirpssgsq
 QY HWSYGLRPGSGSDMSYGLRPGSSOHMSYGLRPGSGSDMSYGLRPGSGSFPKGTAKKI 120
 DB hwsyglirpssgsqdwsgylirpssgsqdwsgylirpssgsqdwsgylirpssgsfpxgaki 120
 QY 121 ILYIPONTYDTFEGNGLDLVKAAEELGIEVOREERNINATQTSIGTQTATGTLTENG 180
 DB 121 Ilyiponydytfeqngldlvkaaeelgievoreerninatqtsigtqtatglteng 180
 QY 181 IYVSAPQIDKLOKTKAGALGSASEIVONANKATVLSGSIOSILSVYAGMLDEALON 240
 DB 181 IYVSAPQIDKLOKTKAGALGSASEIVONANKATVLSGSIOSILSVYAGMLDEALON 240
 QY 241 NSNQHALLARAGLELINSLENIANSVKTLEDFEQISQFSGSKLQNTKGLTGLDKLNIG 300
 DB 241 nsnqhallaqleltnsleniansvktldefeqisqfsgsklqntkglgtldklnig 300
 QY 301 GUDKALGLDVISGLSGATALVLDKKNASTAKKYGAGELANQVYVGNITTKAVSYILA 360
 DB 301 gldkagldgidvisgllsgataalvldkknastakkvgaqfelanqvynitkavssyila 360
 QY 361 ORVAAGLSTGPAALIASTVSLAISPLAFAGIADKFNHAKKESYAERFKKIGYOGDNL 420
 DB 361 orvaaglstgpaalaaistvslaisplafagiadkfnhaksiesyaerfkkigyogdnl 420
 QY 421 LAEYQGTGTIDASVTAINMTALAAGVSAAMAADLTFEKVKNHLYITNSKKEKVTIQMW 480
 DB 421 laeyqgtgtidasvtainmtalaagvsaamaadltfekvknhlyitnskkekvtiqmw 480
 QY 481 FRADAKAEYPNKATKDEKIEIITGONGERTISKOVDLIANGNGKITODELSKYVDNY 540
 DB 481 ftheadakeyppnkatkdekieiitgongertitskovdliangngkitodelskvydny 540
 QY 541 ELTKHKNVTNSDLKLSSVSAFTSSNDSRNVLVAPTSMLOSLSSLOPARSGSOHMSYGL 600
 DB 541 elltkhknvtnsdlklssvsafstssndsrnvlvaptsmldqslssloqfargshmsygl 600
 QY 601 RPSGSGSDMSYGLRPGSSOHMSYGLRPGSGSDMSYGLRPGSSOHMSYGLRPGSGSDM 660
 DB 601 rpsgsqsdmsyglirpssgsqdwsgylirpssgsqdwsgylirpssgsqdwsgylirpssgsqdw 660
 QY 661 SYGLRPGSGSOHMSYGLRPGSGSDMSYGLRPGSGS 695
 DB 661 syglirpssgsqdwsgylirpssgsqdwsgylirpssgs 695

RESULT 2
 Y58361
 ID Y58361 standard; Protein; 695 AA.

XX AC Y58361;
 XX DT 27-MAR-2000 (first entry)
 XX DE Leukotoxin/gonadotropin-releasing hormone fusion protein.
 XX KW Leukotoxin; gonadotropin-releasing hormone; GnRH; immunosterilisation;
 XX KW immunosuppression; vaccine; feline; canine; equine; cervine.
 XX OS Chimeric - Pasteurella haemolytica.
 XX OS Chimeric - Mammalia.
 XX PN WO962545-A2.
 XX PD 09-DEC-1999.
 XX PF 28-MAY-1999; 99WO-CA00493.
 XX PR 04-JUN-1998; 98US-0088024.
 XX PR 06-MAY-1999; 99US-0306689.
 XX PA (BIOS-) BIOSTAR INC.
 XX PI Robbins, SC;
 XX PI MPI; 2000-086857/07.
 XX DR N-PSDB; Z55700.
 PT Hormone immunogens, analogues or antibodies used to manufacture
 PT vaccines for suppression of reproductive behavior and fertility in
 PT vertebrates -
 XX PS Claim 20; Fig 6A-6F; 88bp; English.

CC This sequence represents a fusion protein comprising gonadotropin-
 CC releasing hormone (GnRH) immunogens and a Pasteurella
 CC haemolytica leukotoxin (LKT) protein. The fusion protein comprises, in
 CC the N to C-terminal direction, a synthetic peptide sequence (Y58364), an
 CC eight copy GnRH multimer (composed of two copies of the 4xGnRH multimer
 CC sequence of Y58363), the LKT protein (which functions as a carrier
 CC protein), and a second eight copy GnRH multimer. The fusion protein may
 CC be used in a vaccine composition for prepubertal administration to a
 CC vertebrate subject to result in prolonged suppression of reproductive
 CC behavior and/or fertility. GnRH immunogens, analogues or antibodies
 CC that cross-react with endogenous GnRH of a vertebrate subject are used
 CC to manufacture a composition or vaccine for immunosterilisation or
 CC immunosuppression of feline, canine, equine or cervine subjects.
 CC The vaccines are used to suppress reproductive behaviour and/or
 CC fertility for at least 10 months. The prepubertal administration
 CC results in a prolonged, long-term suppression of testicular development
 CC and/or function in males, or a prolonged, long-term suppression of
 CC ovarian development and/or function in females. The methods provide a
 CC viable and desirable alternative to surgical forms of sterilisation that
 CC are currently used.

XX Sequence 695 AA:

Query Match 100.0%; Score 3591; DB 21; Length 695;
 Best Local Similarity 100.0%; Pred. No. 2.7e-233;
 Matches 695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATVIDRSOHMSYGLRPGSGSDMSYGLRPGSSOHMSYGLRPGSGSDMSYGLRPGSGS
 DB 1 macvdrsqmwsyglirpssgsqdwsgylirpssgsqdwsgylirpssgsqdwsgylirpssgsq
 QY HWSYGLRPGSGSDMSYGLRPGSSOHMSYGLRPGSGSDMSYGLRPGSGSFPKGTAKKI 120
 DB hwsyglirpssgsqdwsgylirpssgsqdwsgylirpssgsqdwsgylirpssgsfpxgaki 120
 QY 121 ILYIPONTYDTFEGNGLDLVKAAEELGIEVOREERNINATQTSIGTQTATGTLTENG 180
 DB 121 Ilyiponydytfeqngldlvkaaeelgievoreerninatqtsigtqtatglteng 180

Db 121 llylpqnyqydeggngldlvkaaeelgievgreemnatatqslgltqraigtterg 180
 Qy 181 IYLSAPOIDKLQKACALGSAESIYONANKAKTVLSGIISIGSVLAGMDLEALON 240
 Db 181 lylsapqldklkqtkagqalgsaeslyvnanakktvlsqslslysvlaagmdlealqn 240
 Qy 241 NSNGHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFSKIQNINKIGTGTGDKLNKNG 300
 Db 241 nsnghalakagletnsleniansvktldefgeqisqfsgklgnkigtltgdklnknig 300
 Qy 301 GLDKAGLGIDVYISGLSGATLALVADKNASTAKKVGAGFELANOVGNITRAVSSYILA 360
 Db 301 gldkaglgldvysglsgataalvldaknaastakkvagfelaangvgnitkavssyila 360
 Qy 361 QRVAGLSSTGVALIISTVSLAISPLAFAGIADKFNNAKLESEYAEFFKLGIDGDL 420
 Db 361 qrvaglsstgvaliistvslaisplafagiadkfnhaksleesyaerfklgydgdl 420
 Qy 421 LAEYRGCTIDASYATINTALAAIAGVSAADLTFEKVHNLVITNSKKEKVTIONW 480
 Db 421 laeyrgctidasvtaintaalaaagvsaadltfekvhnlyitnskkexvtionw 480
 Qy 481 FREADFAKEVPYKATKDEKIEITIGONGERTTSKOVDDLIAGNGKITQDEL SKVVDNY 540
 Db 481 freadfakevpykatkdekieleitigongeritskvoddliakngkitqdel skvvdny 540
 Qy 541 ELKHSKNVTNSLDKLISVSFAFTSSNDSRNVLVAPTSMLODLSLSLOFARQSQHWSTYL 600
 Db 541 ellkshskvtnslldklisvsfaftsndsrnvlvaptsmldslslofarsqhwstyl 600
 Qy 601 RPSGSGQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMS 660
 Db 601 rpsgsgqdwsgylrpgsgsqdwsgylrpgsgsqdwsgylrpgsgsqdwsgylrpgsgsqdw 660
 Qy 661 SYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSS 695
 Db 661 syglrpgsgsqdwsgylrpgsgsqdwsgylrpgsgs 695

RESULT 3
 Y58133
 ID Y58133 standard; Protein: 695 AA.
 XX Y58133;
 AC Y58133;
 XX
 DT 07-MAR-2000 (first entry)
 XX
 De Gonadotropin releasing hormone-leukotoxin fusion protein.
 XX
 KW Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;
 KW antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;
 KW non-androgenic; steroid; reduction; weight gain; muscle distribution;
 KW fat distribution; male pattern; boar taint; flavour; impalement;
 KW reliable; immunocastration; meat production.
 XX
 OS Chimeric - Mammalia.
 OS Chimeric - Pasteurella haemolytica.
 XX
 FN W09956771-A2.
 XX
 PD 11-NOV-1999.
 XX
 PE 05-MAY-1999; 99WO-CA00360.
 XX
 PR 05-MAY-1998; 98US-0084217.
 XX
 PA (BIOS-) BIOSTAR INC.
 XX
 PI Manns JG, Acres SD, Harland R;
 DR WPI: 2000-062125/05.
 DR N-PSDB: 246400.
 XX

PT Production of uncastrated male food animals using vaccines -
 XX
 PS Claim 22; Fig 3A-3F; 87pp; English.
 XX
 CC This sequence represents a chimeric gonadotropin releasing
 CC hormone (GnRH)-leukotoxin (LKT) fusion protein, which may be
 CC used as a vaccine. The LKT portion of the protein acts to enhance
 CC the immunogenicity of the multimeric GnRH portion (Y58133). The invention
 CC relates to a method of using two GnRH immunogen vaccines to produce
 CC uncastrated male animals for meat production, one vaccination prior to
 CC or during the fattening period to reduce circulating testosterone levels,
 CC and the second vaccination about 2-8 weeks before slaughter to
 CC substantially reduce androgenic and/or non-androgenic steroids. The
 CC invention is used to produce food animals that exhibit the weight gain
 CC and muscle/fat distribution of male animals without the problems
 CC associated with male animals. Such problems include "boar taint", a
 CC urine-like odour found in cooked meat of uncastrated pigs which is
 CC caused by steroids stored in the tissues, and similar flavour
 CC impurities in the meat of other intact male animals. The invention is
 CC more reliable than prior art immunocastration techniques.
 CC
 SQ Sequence 695 AA:
 Query Match 100.0%; Score 3591; DB 21; Length 695;
 Best Local Similarity 100.0%; Pred. NO. 2.7e-233;
 Matches 695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MATVIDRSQHWSTYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQ 60
 Db 1 matvidrsgqhwstglrpgsgsqdwsgylrpgsgsqdwsgylrpgsgsqdwsgylrpgsgsq 60
 Qy 61 HWSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSS 120
 Db 61 hwsyglrpgsgsqdwsgylrpgsgsqdwsgylrpgsgsqdwsgylrpgsgsfpkcgakkl 120
 Qy 121 ILYTPONOYDTEGONGIODLVKAAEELGIEVOREERNIATAOISLGTIOAIGLTENG 180
 Db 121 llytpnoydyteggngldlvkaaeelgievgreemnatatqslgltqraigtterg 180
 Qy 181 IYLSAPOIDKLQKACALGSAESIYONANKAKTVLSGIISIGSVLAGMDLEALON 240
 Db 181 lylsapqldklkqtkagqalgsaeslyvnanakktvlsqslslysvlaagmdlealqn 240
 Qy 241 NSNGHALAKAGLELTNSLIENIANSVKTLDEFGEQISQFSKIQNINKIGTGTGDKLNKNG 300
 Db 241 nsnghalakagletnsleniansvktldefgeqisqfsgklgnkigtltgdklnknig 300
 Qy 301 GLDKAGLGIDVYISGLSGATLALVADKNASTAKKVGAGFELANOVGNITRAVSSYILA 360
 Db 301 gldkaglgldvysglsgataalvldaknaastakkvagfelaangvgnitkavssyila 360
 Qy 361 QRVAGLSSTGVALIISTVSLAISPLAFAGIADKFNNAKLESEYAEFFKLGIDGDL 420
 Db 361 qrvaglsstgvaliistvslaisplafagiadkfnhaksleesyaerfklgydgdl 420
 Qy 421 LAEYRGCTIDASYATINTALAAIAGVSAADLTFEKVHNLVITNSKKEKVTIONW 480
 Db 421 laeyrgctidasvtaintaalaaagvsaadltfekvhnlyitnskkexvtionw 480
 Qy 481 FREADFAKEVPYKATKDEKIEITIGONGERTTSKOVDDLIAGNGKITQDEL SKVVDNY 540
 Db 481 freadfakevpykatkdekieleitigongeritskvoddliakngkitqdel skvvdny 540
 Qy 541 ELKHSKNVTNSLDKLISVSFAFTSSNDSRNVLVAPTSMLODLSLSLOFARQSQHWSTYL 600
 Db 541 ellkshskvtnslldklisvsfaftsndsrnvlvaptsmldslslofarsqhwstyl 600
 Qy 601 RPSGSGQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSSQDMS 660
 Db 601 rpsgsgqdwsgylrpgsgsqdwsgylrpgsgsqdwsgylrpgsgsqdwsgylrpgsgsqdw 660
 Qy 661 SYGLRPGSSQDMSYGLRPGSSQDMSYGLRPGSS 695

```

Db      661 syglrpgssqhwsgylrpgsqdwsyglrpgss 695
|||||
RESULT  4
W79570 ID W79570 standard; Protein: 544 AA.
W79570 XX
W79570 XX
24-DEC-1998 (first entry)
XX      LKT-GnRH chimeric protein.
DE      LKT-GnRH chimeric protein.
XX      Chimera; PCB11; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;
KM      cytotoxic activity; antigen presentation; immune response; vaccine;
KM      tumour.
XX
OS      Synthetic.
XX      WO9806848-A1.
XX      PN
XX      PD 19-FEB-1998.
XX      PE 08-AUG-1997; 97WO-CA00559.
XX      PR 09-AUG-1996; 96US-0694865.
XX      PA (UYSA-) UNIV SASKATCHEWAN.
XX      PI Manns JG, Potter AA.
XX      DR WPI; 1998-159540/14.
XX      DR N-PSDB; V61532.
XX      PT Chimeric protein of leukotoxin and gonadotropin releasing hormone
PT      useful for, e.g. preparation of vaccines for reduction of incidence
PT      of mammary tumours in mammals
XX
PS      Disclosure; Figure 7.1-5; 118pp; English.
XX
CC      The present sequence represents the LKT-GnRH chimeric protein from
CC      PCB11. This plasmid contains the LKT 111 polypeptide fused to
CC      four copies of the GnRH peptide. This chimera lacks cytotoxic activity
CC      which enables there to be an increase in antigen presentation and thus an
CC      optimal immune response. The removal of this region also enables the
CC      truncated LKT to be expressed at much higher levels and allows the amount
CC      of antigen administered to be reduced. This chimeric protein comprises a
CC      leukotoxin polypeptide, several multimers, and a GnRH sequence. The
CC      chimeric protein can be used as a vaccine to help reduce the incidence of
CC      mammary tumours in a mammalian individual.
CC
SQ      Sequence 544 AA.

```

```

Query Match      73.6%; Score 2642.5; DB 19; Length 544;
Best Local Similarity 99.4%; Pred. No. 9,6e-170;
Matches 534; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```

```

QY 111 SPFKTAKKIIILYIPONTQYDTGOGNGLODLYAAAEELGIEVORERRNNIATQTSIGTI 170
|||
DB 8 sfpkktakkililyipnyqydtgnglqdlvkaaeelgievgrerinnatqatslgtl 67
QY 171 OFPAIGLTERGIIVSAPOIDKLOKTKAGOLGSAESIYONANKATVLSGIOSILGSYLA 230
|||
DB 68 qtaigttergiivsapqidljqktkagqalgsaesivqnanakakvtl199lgsilgsvla 127
QY 231 GMDLDEALONNSNQHALAKAGLELTNSLIENIANSVKTLDEFEQGIJSQFSKLQNTKIGLG 290
|||
DB 128 gmdldealnnsnqhalakagletlnsleniansvktldefeqglsqfsklnklylg 187
QY 291 TLGDKIKNIGGLDKAGLIGDIVISGLISGATALVLDKNAKSTAKKYGAGFEELANQYVGN 350
|||

```

```

Db      188 lfgdklknigldkaglgldvlsiglsatbaalvjadkhnastakkvagafelanqvgnl 247
QY 351 TRAVSSYIIAQRVAAGLSSGPAVALIASTVSLAISPLAFAGIADKFENHAKSLESYAERP 410
|||
DB 248 tkavssyiliaqrvaaglsstgpaalaaetvslaisplafagladkfnnakslesyaert 307
QY 411 KKLGYDGDMLLAERYRGQTGTIDASVTAINTAALAIAGVSAANA---DLTFEKVKHNLVI 467
|||
DB 308 kklgygdmlaeyrgtgidasvtaintaalaaagvsaaaanlkdlfekkhnlv1 367
QY 468 TNSKKEKVTIQMWERADAFKEVPNFKATKDEKIEELIQNGNRRITSKOYDOLIAAGNCK 527
|||
DB 368 tnskkekvtiqmweradafkevpnfkatkdekielelqngnrritskyvddliakngk 427
QY 528 ITODELSKYVDNTELLKHSKNVTNSLDKLISSVSATSSNDSNNVLVAPTMDSLSST 587
|||
DB 428 itodelskvydnyellkhsknvtnsldkllssvsatssndsnrvlvaptsmdglsst 487
QY 588 QFARGSQHWSYGLRPGSSGSDWSYGLRPGSSQHWSYGLRPGSSQSDWSYGLRPGSS 644
|||
DB 488 qfargsqhwsyglrpgssgdwsgylrpgsqdwsyglrpgsqdwsyglrpgss 544

```

```

RESULT  5
W03943 ID W03943 standard; Protein: 544 AA.
XX
XX      W03943;
AC      W03943;
XX      DT 20-NOV-1996 (first entry)
XX      DE LKT-GnRH protein fusion from PCB11.
XX      KW Leukotoxin; LKT: gonadotropin-releasing hormone; GnRH;
XX      fusion protein; immunogen; vaccine; fertility control;
XX      contraceptive; sterilisation.
XX      OS Chimeric Pasteurella haemolytica A1 strain B122;
XX      OS Chimeric synthetic.
XX
FH      Key
FH      Domain 1..493 Location/Qualifiers
FH      FT Domain /label= LKT
FH      FT 494..544
FH      FT /label= GnRH_repeat_domain
PN      WO9624675-A1.
PD      15-AUG-1996.
XX
XX      24-JAN-1996; 96WO-CA00049.
XX      PF
XX      PR 10-FEB-1995; 95US-0387156.
XX      PA (UYSA-) UNIV SASKATCHEWAN.
XX      PI Manns JG, Potter AA.
XX      DR WPI; 1996-384447/38.
XX      DR N-PSDB; T37177.
XX      PT Gonadotropin-releasing hormone multimer fusion proteins - with
PT      leukotoxin polypeptide for increased immunogenicity, useful in
PT      antifertility vaccine prodn.
XX
PS      Claim 8; Fig 7A-7E; 87pp; English.
XX
XX      A chimeric protein (W03943) is composed of a fusion between
XX      a truncated leukotoxin (LKT-111) from Pasteurella haemolytica
XX      and a 4-copy gonadotropin-releasing hormone (GnRH) repeat
XX      sequence (see also W03944). It is the product of a chimeric
XX      gene (T37177) produced by deleting an approx. 1300 bp sequence
XX      from PCB113 (see also T37176) coding for amino acids 352-784

```

CC of LKT-352. Recombinant plasmid pCB111 (LKT 111:4 copy GnrH,
 CC ATCC 69748) was obt'd. Escherichia coli transformants produced
 CC the chimeric protein, which is useful as a vaccine for fertility
 CC control, esp. immunological sterilisation of domestic or farm
 CC animals.

XX Sequence 544 AA;

Query Match 73.3%; Score 2633.5; DB 17; Length 544;
 Best Local Similarity 99.3%; Pred. No. 3.9e-169;
 Matches 533; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

OY 111 SEPKTGAKKIIILYIPONQYDTEOGNGLODLYKAAEELGIEVOREERNNIATAQTSIGTI 170
 DB 8 sfprtgakkililyipqnydteqngldlykkaeeigievqreernniataqtsigti 67
 OY 171 QTAIGLTERGIVLSAPQIDKLLQTKAGQALGSAESIYQANAKKATVLSGIOSILGSVLA 230
 DB 68 qtaiglttergivilsapqidkllqtkagqalgsaesivqnaankaktvlsigilgsvla 127
 OY 231 GMDLDEALQNNNOHALKAGLELNSLIENANSVKITLDERGEOISQFSGKLQNIKGLG 290
 DB 128 gmdldealqnnnqhalakagletlensliensvktldefeqisqfsgklqnikgig 187
 OY 291 TLGDKLKNIGGLDKAGLDIVISGLSGATAVLADKNASTAKKVGAFELANOVGNI 350
 DB 188 tlgdklknigldkagldivisgllsgataalvladknaastakkvagfelaanvgvni 247
 OY 351 TKAVSSYTLAORVAAGLSSTGFPVALLASTVSLAISPLAFAGIADKFNHAKSLESYAERF 410
 DB 248 tkavssyllaorvaaaglsstgfpvaallastvslaisplafagiadkfnhakslesyaerf 307
 OY 411 KKLGYDGNLAEYRGCTTIDASTAINTALAIAGVSAAA--DLTEPKVKNHNLVI 467
 DB 308 kklgydgnllaeyrgctidaastaintalaaagvsaaanldltefkvknhnlvi 367
 OY 468 TNSKKEKTIQWMPREAPFAKPEVPYKATKDEKIEIIGONGERTSKROYDILKKNK 527
 DB 368 tnskkekvtiqwmpreatfakpevpnykatkdekieiigongertskroydilkknk 427
 OY 528 ITODELSKRVNDELKSKNVTNSLDKLSSVSASTSSNDNRNVLPVPTSMLDQSLSL 587
 DB 428 itodelskrvndellksknvtnslldklssvsastssndnrnlvptsmldqslsl 487
 OY 588 OFARSSQHWSTGLRPGSSQDMSTGLRPGSSQDMSTGLRPGSSQDMSTGLRPGSS 644
 DB 488 ofarssqhwstglrpgssqdmstglrpgssqdmstglrpgssqdmstglrpgss 544

RESULT 6
 ID W03942 standard: Protein; 977 AA.

XX W03942;

XX 20-NOV-1996 (first entry)

XX LKT-GnrH protein fusion from pCB113.

XX Leukotoxin; LKT; gonadotropin-releasing hormone; GnrH;
 KW fusion protein; immunogen; vaccine; fertility control;
 KW contraceptive; sterilisation.

XX Chimeric Pasteurella haemolytica A1 strain B122;
 OS Chimeric synthetic.

XX Key location/Qualifiers

XX 1.929

XX /label= LKT

XX Domain 927..977

XX /label= GnrH_repeat_domain

PN W09624675-A1.
 XX 15-AUG-1996.
 PD 24-JAN-1996; 96WO-CA00049.
 XX 10-FEB-1995; 95US-0387156.
 XX (UUSA-) UNIV SASKATCHEWAN.
 XX Manns JG, Potter AA;
 XX WPI: 1996-384447/38.
 DR N-PSDB; T37176.

PT Gonadotropin-releasing hormone multimer fusion proteins - with
 PT leukotoxin polypeptide for increased immunogenicity, useful in
 PT antifertility vaccine prodn.

PS Claim 7; Fig 5A-5H; 87pp; English.

CC A chimeric protein (W03942) is composed of a fusion between
 CC a truncated Leukotoxin (LKT-352) from Pasteurella haemolytica (see
 CC also W03945) and a 4-copy gonadotrophin-releasing hormone (GnrH)
 CC repeat sequence (see also W03944). It is the product of a
 CC chimeric gene (T37176) produced by ligating a synthetic sequence
 CC for the 4-copy GnrH into vector pAA352 (ATCC 68283), which carries
 CC the LKT-352 gene. Recombinant plasmid pCB113 (LKT 352:4 copy
 CC GnrH, ATCC 69749) was obt'd. Escherichia coli transformants
 CC produced the chimeric protein, which is useful as a vaccine for
 CC fertility control, esp. immunological sterilisation of
 CC domestic or farm animals.

XX Sequence 977 AA;

Query Match 67.6%; Score 2426; DB 17; Length 977;
 Best Local Similarity 55.1%; Pred. No. 7.5e-155;
 Matches 534; Conservative 0; Mismatches 0; Indels 436; Gaps

OY 111 SEPKTGAKKIIILYIPONQYDTEOGNGLODLYKAAEELGIEVOREERNNIATAQTSIGTI 170
 DB 8 sfprtgakkililyipqnydteqngldlykkaeeigievqreernniataqtsigti 67
 OY 171 QTAIGLTERGIVLSAPQIDKLLQTKAGQALGSAESIYQANAKKATVLSGIOSILGSVLA 230
 DB 68 qtaiglttergivilsapqidkllqtkagqalgsaesivqnaankaktvlsigilgsvla 127
 OY 231 GMDLDEALQNNNOHALKAGLELNSLIENANSVKITLDERGEOISQFSGKLQNIKGLG 290
 DB 128 gmdldealqnnnqhalakagletlensliensvktldefeqisqfsgklqnikgig 187
 OY 291 TLGDKLKNIGGLDKAGLDIVISGLSGATAVLADKNASTAKKVGAFELANOVGNI 350
 DB 188 tlgdklknigldkagldivisgllsgataalvladknaastakkvagfelaanvgvni 247
 OY 351 TKAVSSYTLAORVAAGLSSTGFPVALLASTVSLAISPLAFAGIADKFNHAKSLESYAERF 410
 DB 248 tkavssyllaorvaaaglsstgfpvaallastvslaisplafagiadkfnhakslesyaerf 307
 OY 411 KKLGYDGNLAEYRGCTTIDASTAINTALAIAGVSAAA----- 454
 DB 308 kklgydgnllaeyrgctidaastaintalaaagvsaaagvsiaspiallvsgit 367
 OY 455 ----- 454
 DB 368 gvisitlgyskqamfehanklnkiveweknbnhkyfengydarylanlqdmkfln 427
 OY 455 ----- 454
 DB 428 lnkelaervalitqgwdnlnigdlagisrlgekvlsgkayvdafegkhkacklvgld 487
 OY 455 ----- 454

Db 488 sanglidsnsgkaktqhlfrp1ltpgrehervqtkyeyltklnrvswkltdg 547
 QY 455 ----- 454
 Db 548 aasftldlnvgrlgeldnagnvltkktkllaklgdgnvfvgsgteldggyd 607
 QY 455 ----- 454
 Db 608 rvhysrgnygaltdatketegsgyfvnrfvetgkahlvsthtalvgnreekieyrhs 667
 QY 455 ----- 454
 Db 668 nmgahgyytkdtklkaeeiigtshndifkyskfndafnggdvdtldgndnrlfgyk 727
 QY 455 ----- 454
 Db 728 gddldgngddfidggkngdlhhgkgddlfvhrkygdndltdsgndklfsdsnlk 787
 QY 455 DLFEKVKHNLVTNNSKKEKVTIQQNMFREADFAKEVNPYKATDEKIEITIGONGERITS 514
 Db 788 dlfekvkhnvltnskkekvltgnwtfreadfakvnpnykatdekleeiignggerits 847
 QY 515 KQYDDLIAKNGKITODELSKVVDNTELLKHSKNVTNSLDKLSSVSAFTSSNDNRNVLV 574
 Db 848 kyvddliakngkltqdeliskvvdnyellkhsknvtnsldklssvsaftssndsrnvlv 907
 QY 575 APTSMIDQSSIOFARGSOHWSYGLRPGSGSDMSYGLRPGSSOHWYGLRPGSGSD 634
 Db 908 apsmidqssisqfargsqhwsyglrpgssqdsyglrpgssqhsyglrpgssqds 967
 QY 635 WSYGLRPGGS 644
 Db 968 wsyglrpggs 977
 RESULT 7
 ID W79569 standard; Protein; 977 AA.
 AC W79569;
 DT 24-DEC-1998 (first entry)
 DE LKT-GnRH chimeric protein.
 KW Chimera; pcB113; LKT 352; GnRH; Gonadotropin releasing hormone; multimer;
 KW cytotoxic activity; antigen presentation; immune response; vaccine;
 KW tumour.
 OS Synthetic.
 PN WO9806848-A1.
 PD 19-FEB-1998.
 PF 08-AUG-1997; 97MO-CA00559.
 PR 09-AUG-1996; 96US-0694865.
 PA (UYSA-) UNIV SASKATCHEWAN.
 PI Manns JG, Potter AA;
 DR WPI; 1998-159540/14.
 DR N-PSDB; V61531.
 PT Chimeric protein of leukotoxin and gonadotropin releasing hormone
 PT useful for, e.g. preparation of vaccines for reduction of incidence
 PT of mammary tumours in mammals
 XX
 PS Disclosure; Figure 5.1-8; 118pp; English.

CC The present sequence represents the LKT-GnRH chimeric protein from
 CC pcB113. This plasmid contains the LKT 352 polypeptide (W79568) fused to
 CC four copies of the GnRH peptide. This chimera lacks cytotoxic activity
 CC which enables there to be an increase in antigen presentation and thus an
 CC optimal immune response. The removal of this region also enables the
 CC truncated LKT to be expressed at much higher levels and allows the amount
 CC of antigen administered to be reduced. This chimeric protein comprises a
 CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The
 CC chimeric protein can be used as a vaccine to help reduce the incidence of
 CC mammary tumours in a mammalian individual.
 XX
 SQ Sequence 977 AA;
 Query Match 67.6%; Score 2426; DB 19; Length 977;
 Best Local Similarity 55.1%; Pred. No. 7.5e-155;
 Matches 534; Conservative 0; Mismatches 0; Indels 436; Gaps 1;
 QY 111 SPFKTGAKKILLYIPONYOYTFEOGNGLODLVKAABELGIEVOREBRNNIATQTSIGTI 170
 Db 8 sfktgakkillyipnnyqdtgnglqdlvkaaeelgievgreennlatatqstlgtl 67
 QY 171 QTAIGTREGIVYSAPOIDKLLKTKAGALGSABESIVONANKAKTVLSGTSISVLA 230
 Db 68 qtaigtregivisapqidlkltkagagalsaesivgnankakvtslgisllsvla 127
 QY 231 GMDLDEALONNSNOHALAKGLELTNSLIENANSVKTIDFEFEOISOPFSKLONTKGLG 290
 Db 128 gmdldealgmsnqhatakgletnslienansvktidefegqisqfysknglkyig 187
 QY 291 TLGDKLKNIGGLDKAGIGLIDVISGLSGATALVLDKRNASTAKKYGAGELANOVVNT 350
 Db 188 tlgdklknlgldkagigldvisgllsgataalvldknaastakkyagfelanovynt 247
 QY 351 TKAVSSYTLAORVAAGSSGPAALIASVSLASPLAAGTADFNHAKSLESYAEPR 410
 Db 248 tkavssytlagrvaagssgpaaliasvslasplataagtkfnhakslesyaerf 307
 QY 411 KRLGDDNLLAEYORGTGTIDASVTAINPALAIGVSAANA----- 454
 Db 308 krlgddnllaeqrgtgtldasvtaintalaigvsaanaagsviallvsqilt 367
 QY 455 ----- 454
 Db 368 gvlstllqysqgmfehvanklnhkiwelnhgnknyfengydarylanlgdmkflln 427
 QY 455 ----- 454
 Db 428 lnkelaervialtqgqwdnlgdlaglsrlgkvlsgkayvdafeegkhikadklvqld 487
 QY 455 ----- 454
 Db 488 sanglidsnsgkaktqhlfrp1ltpgrehervqtkyeyltklnrvswkltdg 547
 QY 455 ----- 454
 Db 548 aasftldlnvgrlgeldnagnvltkktkllaklgdgnvfvgsgteldggyd 607
 QY 455 ----- 454
 Db 608 rvhysrgnygaltdatketegsgyfvnrfvetgkahlvsthtalvgnreekieyrhs 667
 QY 455 ----- 454
 Db 668 nmgahgyytkdtklkaeeiigtshndifkyskfndafnggdvdtldgndnrlfgyk 727
 QY 455 ----- 454
 Db 728 gddldgngddfidggkngdlhhgkgddlfvhrkygdndltdsgndklfsdsnlk 787
 QY 455 DLFEKVKHNLVTNNSKKEKVTIQQNMFREADFAKEVNPYKATDEKIEITIGONGERITS 514
 Db 788 dlfekvkhnvltnskkekvltgnwtfreadfakvnpnykatdekleeiignggerits 847

QY 515 KOVDLAKNGKTIPTODELSKVNDVELLKHSKNVTNSLDKLSSVAFSTSSNDSRNLY 574
 DB 848 KQVDDLLAKNGKTIPTODELSKVNDVELLKHSKNVTNSLDKLSSVAFSTSSNDSRNLY 907
 QY 575 APTSMLOSLSLQFANGSQHMSYGLRPGSGSDMSYGLRPGSSQHMSYGLRPGSGSD 634
 DB 908 APTSMLOSLSLQFANGSQHMSYGLRPGSGSDMSYGLRPGSSQHMSYGLRPGSGSD 967
 QY 635 WSYGLRPGSG 644
 DB 968 WSYGLRPGSG 977

RESULT 8
 Y33929
 ID Y33929 standard; Protein; 490 AA.
 XX
 AC Y33929;
 XX
 DT 09-NOV-1999 (first entry)
 XX
 DE Leukotoxin carrier protein (LKT 114).
 XX
 KM Myostatin: mouse; rabbit; human; baboon; bovine; porcine; ovine; chick;
 KM turkey; zebrafish; immune response; vaccine; body weight; muscle mass;
 KM mammary gland tissue; lactation; feed uptake; muscle degeneration;
 KM GDF11 activity; LKT 114; leukotoxin carrier protein.
 XX
 OS Pasteurella haemolytica.
 XX
 PN WO9942573-A1.
 XX
 PD 26-AUG-1999.
 XX
 PF 19-FEB-1999; 99WO-CA00128.
 XX
 PR 19-FEB-1998; 98US-0075213.
 XX
 PA (BIOS-) BIOSTAR INC.
 XX
 PI Barker CA, Morsey M;
 XX
 DR WPI: 1999-527471/44.
 DR N-PSDB: X99361.
 XX
 PT New myostatin peptide, multimers and immunocjugates for eliciting
 XX an immune response in a vertebrate against a myostatin immunogen
 XX
 PS Example 2; Fig 15; 109pp; English.
 XX
 CC The invention provides myostatin peptides consisting of 3-100 amino
 CC acids, derived from a region of mouse, rabbit, human, baboon, bovine,
 CC porcine, ovine, chick, turkey or zebrafish myostatin (see sequences
 CC Y33930-939). The myostatin peptides are derived preferably from a region
 CC of amino acid residues 1-275, 25-300, 50-325 or 75-350 of the above
 CC sequences. The peptides and the nucleic acids encoding the peptides are
 CC useful as vaccines for eliciting an immune response in a vertebrate
 CC against a myostatin immunogen. They result in increasing body weight,
 CC muscle mass, number and size of muscle cells, muscle strength, mammary
 CC gland tissue, lactation, appetite or feed uptake, life span of the
 CC vertebrate, and cause a reduction in body fat content, useful for muscle
 CC wasting conditions. The vaccines are also useful for treating a disorder
 CC which comprises degeneration or wasting of muscle in a vertebrate, and
 CC useful for modulating GDF11 activity. The present sequence represents a
 CC leukotoxin carrier protein (LKT 114) which is used as an immunological
 CC carrier protein to form a myostatin immunocjugate.
 XX
 SQ Sequence 490 AA;

Query Match 65.6%; Score 2354; DB 20; Length 490;
 Best Local Similarity 100.0%; Pred. No. 2e-150;

Matches 483; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 111 SFPKTKAKKIIITIPQNTQYDTGQNGLODYKAAEELGIEVQREERNNTAQTSLGTI 170
 DB 8 SFPKTKAKKIIITIPQNTQYDTGQNGLODYKAAEELGIEVQREERNNTAQTSLGTI 67
 QY 171 QTAIGLTERGIVLSAPQIDKLOKTKAQALGSAESIYQNNKAKTVLSGIOSIIGSLA 230
 DB 68 QTAIGLTERGIVLSAPQIDKLOKTKAQALGSAESIYQNNKAKTVLSGIOSIIGSLA 127
 QY 231 GMDLDEALQNNNSNOHALAKAGLELTNSLIENIANSVKTLDERFGEQISQSKLQNKIGL 290
 DB 128 GMDLDEALQNNNSNOHALAKAGLELTNSLIENIANSVKTLDERFGEQISQSKLQNKIGL 187
 QY 291 TLGDKLKNIGGLDKRAGLGVISGLSGATPAAVLADKNASTAKKVGAGFELANQVGN 350
 DB 188 TLGDKLKNIGGLDKRAGLGVISGLSGATPAAVLADKNASTAKKVGAGFELANQVGN 247
 QY 351 TKAVSSYTLAQRVAAAGLSSTGPVAALIASTVSIATSPAFGIADKFNHAKSLESTAE 410
 DB 248 TKAVSSYTLAQRVAAAGLSSTGPVAALIASTVSIATSPAFGIADKFNHAKSLESTAE 307
 QY 411 KKLGYDGNLAEYQRTGTIDASVTATNTALAAIAGCVSAAADLPFEKYKHNLYTNS 470
 DB 308 KKLGYDGNLAEYQRTGTIDASVTATNTALAAIAGCVSAAADLPFEKYKHNLYTNS 367
 QY 471 KKEKYTTNNMFREADFAKEVPNYKATKDEKTEBIIIGNGERTSKQVDLAKNGKTIQ 530
 DB 368 KKEKYTTNNMFREADFAKEVPNYKATKDEKTEBIIIGNGERTSKQVDLAKNGKTIQ 427
 QY 531 DELSKVVDNVELLKHSKNVTNSLDKLSSVAFSTSSNDSRNLYAPTSMLOSLSLQFA 590
 DB 428 DELSKVVDNVELLKHSKNVTNSLDKLSSVAFSTSSNDSRNLYAPTSMLOSLSLQFA 487
 QY 591 RGS 593
 DB 488 RGS 490

RESULT 9
 R34547
 ID R34547 standard; Protein; 936 AA.
 XX
 AC R34547;
 XX
 DT 23-AUG-1993 (first entry)
 XX
 DE GNRH-leukotoxin gene fusion prod.
 XX
 KM Vector; LKT 352; flanking; recombinant; antigen; somatostatin;
 KM gonadotropin releasing hormone; rotavirus viral protein 4;
 KW carrier protein; lactation; reproduction; SRIF.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Protein 1..926
 FT Peptide /note="recombinant leukotoxin protein"
 FT 927..936
 FT /note="GNRH"
 XX
 PN WO9308290-A.
 XX
 PD 29-APR-1993.
 XX
 PF 15-OCT-1992; 92WO-CA00449.
 XX
 PR 16-OCT-1991; 91US-0779171.
 PR 14-OCT-1992; 92US-0960932.
 XX
 PA (UYSA-) UNIV SASKATCHEWAN.
 XX
 PI Hughes HPA, Potter AA, Redmond MJ;

XX WPI: 1993-152482/18.
 DR N-PSDB: Q41322.
 XX
 PT Immunological carrier system with enhanced immunogenicity -
 PT comprises chimeric protein comprising leuco:toxin peptide or
 PT homologous protein fused to antigen esp. somatostatin or
 PT gonadotropin releasing hormone
 XX
 PS Example 2; Fig 8; 95pp; English.
 XX
 CC Oligonucleotides contg. sequences from bovine gonadotrophin
 CC releasing hormone (GnRH) gene were constructed on a Pharmacia Gene
 CC Assembler using standard phosphoramidite chemistry. The oligo-
 CC nucleotides were annealed and ligated into vector pAA352 (contg.
 CC the Pateurella haemolytica leuco-toxin gene) which had been digested
 CC with BamHI. The ligated DNA was used to transform E. coli strain
 CC MH3000. Transformants contg. the oligonucleotide inserts were
 CC identified by restriction endonuclease mapping and the recombinant
 CC plasmid designated pAA502. The chimeric protein produced from the
 CC plasmid works to bring about a larger immune response than the antigen
 CC alone, i.e. the leukotoxin works as a carrier protein.
 CC See also R34545-8.
 CC
 SQ Sequence 936 AA:

Query Match 60.5%; Score 2173; DB 14; Length 936;
 Best Local Similarity 52.9%; Pred. No. 6,9e-138;
 Matches 491; Conservative 0; Mismatches 2; Indels 436; Gaps 1;

QY 111 SFRPTGAKIILLYPONYQYDEQNGLODLVRAEELGIEVORENNNTATQTSIGTI 170
 |||||||
 Db 8 sftkgakakillipnyqydeqngldlvkaaeelgvegreenmataqtstglt 67
 |||||||
 QY 171 QTAIGTERTGIVSAPQIDKLOKTRAGALGSAESTIVONANKAKTYLSGISISVLA 230
 |||||||
 Db 68 qtaigttergivilsapqidlqktkagalgsgaestivgnankaktylsqisllgsvla 127
 |||||||
 QY 231 GMLIDEALONNSQNHAKRAGELTNSLIENINSVKTIDEFEQIOISQESKIONKIG 290
 |||||||
 Db 128 gmlidealqunsnghalakageltnslleiansvktlidxceqlsqkignkylg 187
 |||||||
 QY 291 TLGDCKRNITGDLKAGIGLDIVISGLSGATPAALVLDKNASTAKKYGAGFELANQVYVNT 350
 |||||||
 Db 188 tlgdcklknigldkaglgldivisgllsgataalvldknaastakkygagfelanqyvgnl 247
 |||||||
 QY 351 TKAVSSYIIAQRVAAGLSTGPVPAALIASTVSAISPAPAGIADKFNNAKSLSTAEERF 410
 |||||||
 Db 248 tkavssyilagrvaaglstgypvaalvastvsaistplafadkfhnakslesyerf 307
 |||||||
 QY 411 KKLGYGDNLAEYRGSTGTIDASVTAINTALAIAGVSAANA----- 454
 |||||||
 Db 308 KKLgygdnlleayrgrtgcldasvtalntalaalegvsaaagvsiasplallvsqit 367
 |||||||
 QY 455 ----- 454
 Db 368 gvtstllqyskqamfehvancklnkiveweknbnhknfyfengydarylanlqdmkflin 427
 |||||||
 QY 455 ----- 454
 Db 428 lnkelaervialtggqwdnlgdlaglsrlgskvlskayvdafeegkhikadklvql 487
 |||||||
 QY 455 ----- 454
 Db 488 sanglidsvnsqakcqhllfrtptlltpgtehrerwqtgkyeytklnlnrvdswkilt 547
 |||||||
 QY 455 ----- 454
 Db 548 aasstfdlnvgrjgldnagnvktketkiaklgdgdanvfgsgteldggyd 607
 |||||||
 QY 455 ----- 454

Db 608 rvhsrgnygaltidakebegsyfwnrfvelgkalhvevsthlalvgnreekieyrhs 667
 |||||||
 QY 455 ----- 454
 Db 668 nngnhagyytkdtkaveeilgtshndifksgkfndafnggdvdtidngdnndrlfsgk 727
 |||||||
 QY 455 ----- 454
 Db 728 gddldgngddfidgkgndllhggkgddlfvhrkgdgnlldsdgndklsfsdsnlk 787
 |||||||
 QY 455 DLTFEKVKHNLVITNSKKEVYTIQNMFRADPAKEVPNPKATYDEKIEELIGONGERITS 514
 |||||||
 Db 788 dltfekvkhnlvltnskkekvltlgnwfreaddakevnykatkdekleeilgngerits 847
 |||||||
 QY 515 KOYDDLIAGKNGKITODELSKVVDNELKHSKNVTNSIDKLSSVSAPFSSNDSRNVLV 574
 |||||||
 Db 848 kyvddliakngkitqdeliskvvdnyellkhsnvnnsldklssvsafssndsrnvlv 907
 |||||||
 QY 575 APTSMLOSLSSLOFARGSQHWSYGLRPG 603
 |||||||
 Db 908 aptsmldgslsslqfargsqhwsyglrpg 936
 |||||||
 RESULT 10
 R52748
 ID R52748 standard; Protein; 1069 AA.
 XX
 AC R52748;
 XX
 DT 01-JUL-1994 (first entry)
 XX
 DE Bovine IFNgamma/LTK chimeric protein encoded by plasmid pAA497.
 XX
 DE Bovine: Interleukin-2; IL2; P. haemolytica; leukotoxin; LTK; IFN.
 KW ItkA; chromosome walking; fusion protein; vaccine; interferon; gamma;
 KW monoclonal; polyclonal; antibody.
 XX
 OS Pasteurella haemolytica - chimera.
 OS Bos taurus - chimera.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..926
 FT /note= "Recombinant leukotoxin peptide [splic]"
 FT Peptide 927..1069
 FT /note= "Bovine IFNgamma"
 XX
 PN US5273889-A.
 XX
 PD 28-DEC-1993.
 XX
 PF 22-AUG-1990; 90US-0571301.
 XX
 PR 22-AUG-1990; 90US-0571301.
 PR 16-OCT-1991; 91US-0777715.
 XX
 PA (CIBA) CIBA GEIGY CANADA LTD.
 PA (UYSA-) UNIV SASKATCHEWAN.
 XX
 PI Campos M, Hughes HPA, Potter A;
 XX
 DR WPI: 1994-006687/01.
 DR N-PSDB: Q54213.
 XX
 PT Immunogenic fusion proteins of gamma-interferon and immunogenic
 PT leukotoxin - used in vaccines and to raise monoclonal and polyclonal
 PT antibodies
 XX
 PS Disclosure; Fig 7; 56pp; English.
 XX
 CC This sequence represents a fusion between bovine gamma interferon
 CC (IFNgamma) and Pasteurella haemolytica leukotoxin (LTK). The
 CC leukotoxin gene, ItkA, was isolated from a gene library of P.
 CC haemolytica by chromosome walking. Immunogenic fusion proteins

CC such as this can be used in vaccine compositions. It can also be
 CC used to raise mono- and polyclonal antibodies.

xx Sequence 1069 AA;

Query Match 59.3%; Score 2131; DB 15; Length 1069;
 Best Local Similarity 52.6%; Pred. No. 5.5e-135;
 Matches 484; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

```

OY 111 SFPKGAKKIILYIPONTQYDTEGNGLDLVKAAEELGIEVQREERNINATQTSIGTI 170
DB 8 sfptgakkililyipqnygdtdegngldlvkaaeelgievqreerninatqtslgti 67
OY 171 QTAITLTERGTYLSAPQIDKLLQKTKAGALGSAESTYQVNAKAKTVLSIGTISIGSVLA 230
DB 68 qtaigttergtylsapqidkllqktkagqalsaesivqnaekaktvlsigtsilgsvla 127
OY 231 GMDLDEALQNNNSQHALKAGLELTNSLTENANSVKTTDEFGEOISQFSKLNKIGLG 290
DB 128 gmdldealqnnnsqhalakagletnsltenansvktldefgesqfsklnqtkgig 187
OY 291 TLGDKLNKINGIDKAGLGIDVISGLSGATAALVLADKNASTAKRVGAGFELANOVVNI 350
DB 188 tlgdklnkingidkaglgidvisgllsgataalvladknastakrvagagfelanqvgni 247
OY 351 TKAVSSYLLAQRVAVAGLSTGPAVALLASTVSLAISPFAIADKFNHAKLESTAEAF 410
DB 248 tkavssyllaqrvavaglstgpvaallastvslasplafaiadkfnhakslesyaerf 307
OY 411 KKLGYDGNLLAEYORGTIDASTATNTALATAGVSAANA----- 454
DB 308 kklgydgnllaeaygrgtidastratntalaalagvsaanaagvsplallvsgit 367
OY 455 ----- 454
DB 368 gvlstllgysqamfehvaanklhkiveweknbgknyfengydarylanlqdmkflin 427
OY 455 ----- 454
DB 428 lnkelaervlaiaiqgwdnlnlgiagisrlgekvlsqayvdafeegkhikadlvqld 487
OY 455 ----- 454
DB 488 sanglidsnsgaktqhlfrtplltpgtetrrervgqkyeyicklninrvdswkicldg 547
OY 455 ----- 454
DB 548 aasstfdltvnvqrigeldnagvntckekliaklgeddnvfvsgtteidsggyd 607
OY 455 ----- 454
DB 608 rvhvsrgnygaltdacketeqsgyltnrfvetgkahlhevtsthalvgnreekleyrth 667
OY 455 ----- 454
DB 668 nqnhaqyrtkdtlkaveeilgtsmndifkgsfndafrngddgvdtidgndndlfggk 727
OY 455 ----- 454
DB 728 gddlldgngddfldgqgndllbhgkgddlfvhrkgdgnndltdsdgndklsfedsnlk 787
OY 455 ----- 454
DB 788 dltefkvhnvltmskkekvtlqmwfpreadpapakvnpnykatkdkiieitigoneerits 514
DB 847 dltefkvhnvltmskkekvtlqmwfpreadpapakvnpnykatkdkiieitigoneerits 847
OY 515 KOVDLIAKNGNKITQDELSEKVVNDVEYELLSKNTNSLDELKISSVSAFTSSNDERNVLY 574
DB 848 kqvddlakngnkltqdelsekvvndveyellshkntnsldeklissvsaftssndsrnvly 907
OY 575 APTSMLOSLSLOPARCSQ 594
DB 908 aptsmldsqslsifargsq 927

```

```

RESULT 11
ID W13867
XX W13867 standard; Protein; 1069 AA.
AC W13867;
XX 12-MAR-1997 (first entry)
DT 12-MAR-1997 (first entry)
DE Chimeric protein #2.
XX
XX RTX cytotoxin; cytokine; immunogen; chimeric protein; cytokine; vaccine;
KW interleukin-2, IL-2; gamma interferon; gamma IFN; leukotoxin; pneumonia;
KW Pasteurella haemolytica; LKT352; respiratory disease; shipping fever;
KW fibrinous pneumonia; cattle; therapy.
XX
OS Synthetic.
XX
XX US5594107-A.
XX
XX 14-JAN-1997.
XX
XX 22-AUG-1990; 90US-0571301.
XX
XX 20-DEC-1993; 93US-0170126.
XX
XX 22-AUG-1990; 90US-0571301.
XX
XX 16-OCT-1991; 91US-0777715.
XX
PA (CIBA ) CIBA GEIGY CANADA LTD.
PA (UYSA-) UNITV SASKATCHEWAN.
XX
PI Campos M, Hughes HPA, Potter A;
XX
XX WPI; 1997-099529/09.
XX
DR N-PSDB; T60033.
XX
XX Immunogenic chimeric proteins comprising cytokine linked to RTX
PT toxin - useful in vaccines, esp. against shipping fever in cattle
XX
XX Claim 13; Column 37-46; 56pp; English.
XX
XX W13866 and W13867 represent immunogenic chimeric proteins of the
CC invention. This sequence represents a chimeric protein containing the
CC bovine gamma interferon (gamma IFN) sequence and a leukotoxin sequence.
CC The chimeric proteins of the invention comprise a cytokine, selected from
CC interleukin-2 (IL-2) and gamma IFN, linked to at least one RTX toxin
CC epitope (preferably the sequence shown in W13865). The RTX toxin used to
CC provide the epitope sequence is preferably a leukotoxin, especially the
CC full-length Pasteurella haemolytica leukotoxin. Alternatively, the
CC leukotoxin is a truncated leukotoxin lacking leukotoxic activity,
CC especially LKT352. The chimeric proteins can be used for the production
CC of vaccines against respiratory diseases such as pneumonia, particularly
CC fibrinous pneumonia caused by P. haemolytica, including shipping fever in
CC cattle.
XX
XX
SO Sequence 1069 AA;

Query Match 59.3%; Score 2131; DB 18; Length 1069;
Best Local Similarity 52.6%; Pred. No. 5.5e-135;
Matches 484; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

OY 111 SFPKGAKKIILYIPONTQYDTEGNGLDLVKAAEELGIEVQREERNINATQTSIGTI 170
DB 8 sfptgakkililyipqnygdtdegngldlvkaaeelgievqreerninatqtslgti 67
OY 171 QTAITLTERGTYLSAPQIDKLLQKTKAGALGSAESTYQVNAKAKTVLSIGTISIGSVLA 230
DB 68 qtaigttergtylsapqidkllqktkagqalsaesivqnaekaktvlsigtsilgsvla 127
OY 231 GMDLDEALQNNNSQHALKAGLELTNSLTENANSVKTTDEFGEOISQFSKLNKIGLG 290

```

```

Db 128 gmdldeaiqnnsnqhalakagletnsljenlansvkldefeqsgfysklnkylg 187
QY 291 TLGDKLKNIGLDKAGLGLDVISGLSGATPAALVLADKNASTKRVGAGFELANQVYVNT 350
Db 188 tlgdtklnlgldkaglgldvlsqllsgataalvldknaastakvvgagfelanqvyvnt 247
QY 351 TKAVSSYIIAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNNAKSLESAERF 410
Db 248 tkavssyllagrvaaglsstgpvaalaaastvslaispladkfnhakslesyaerf 307
QY 411 KRLGYGDNILAEYQRTGTIDASVTAINATALAINGVSAAAA----- 454
Db 308 krlgygdnllaeayrgrgtldasvtaintaalaagvsaaagvsiasplallvsglt 367
QY 455 ----- 454
Db 368 gvisltllgyskgamfehvanklnkiveweknmhknfyfengydarylanldnmkflin 427
QY 455 ----- 454
Db 428 lnkelgaerviaitqgwdnigldaglsrlgkvisgkayvdafeegkhikadkivld 487
QY 455 ----- 454
Db 488 saanglidsvnsqaktqhlffrtplltpgtbhervgtkyeylklalnrvdswkiltg 547
QY 455 ----- 454
Db 548 aaastfdlcnvgrislaidnagnvktketkliakligegdanvfvgstleldgegyd 607
QY 455 ----- 454
Db 608 rvhysrgnygaltidatketegsylvnrfvetgkalhevtsthalvgnreekleyrha 667
QY 455 ----- 454
Db 668 nqghagyytkdtkaveeilgtshndlfkyskfndafngdgvdtdldgndgnrlfgyk 727
QY 455 ----- 454
Db 728 gddilgngdgdldidgqkgnldlhngkgyddlfvhrkygdnditdsgdnklfsdsnlk 787
QY 455 DLTFEKVKNLVTNKKKEKVTIONMFREADFAKEVPNKATFDEKIEELIIGONGERITS 514
Db 788 dltfekvknlvitnkkkvtlqnfreadfakvpenykatdekieleligngerits 847
QY 515 KQYVDLIANGNGKITODELSKVYDNEELKHSKNVNTSLDKLISSVSATFSSWDSRWLV 574
Db 848 kqvddliakngkitqdeliskvdyneellkhsknvntsldkliassvsatfssndsrnvlv 907
QY 575 APTSMIDQSLSSLOFARGSQ 594
Db 908 aptsmldqslsslqfargsq 927

```

```

RESULT 12
R14482
ID R14482 standard; Protein: 926 AA.
XX
AC R14482;

```

```

XX 15-JAN-1992 (first entry)
XX
DE LKT352.
XX
KW Antigen; leukotoxin; vaccine; Ikta.
XX
OS Pasteurella haemolytica.
XX
PN WO9115237-A.
XX
PD 17-OCT-1991.
XX

```

```

PF 17-OCT-1991; 91WO-CA00170.
XX
PR 05-APR-1990; 90US-0504850.
XX
PA (UYSA-) UNIV SASKATCHEWAN.
XX
PI Acres SD, Barluk LA, Potter AA, Lawman MDP;
DR WPI; 1991-324967/44.
XX
PT Vaccines for Pasteurella haemolytica infection in cattle -
PT comprise sub-unit antigens from P haemolytica fimbrial protein,
PT plasmid receptor, 50 K outer membrane protein and leukotoxin.
XX
PS Disclosure; Fig 5; 92pp; English.
XX
CC LKT352 is 98% homologous with authentic leukotoxin and migrates
CC to the same position on gels.
CC The LKT352 gene was prep'd. as follows: Ikta, an MaeI fragment
CC contg. the gene was ligated into the SmaI site of pUC13 to form
CC pAA179. From this, two constructs were made in the prac-based
CC vector, pGH432:laci digested with SmaI. One, pAA342, consisted of
CC the 5' AhairII fragment from Ikta while the other, pAA345, contained
CC the entire MaeI fragment. Clone pAA342 expressed a truncated
CC leukotoxin peptide at high levels while pAA345 expressed full
CC length leukotoxin at very low levels. The 3' end of the Ikta gene
CC of pAA345 was therefore ligated to StyI/BamHI digested pAA342 to
CC yield pAA352 contg. the LKT352 sequence. The protein expressed
CC from the vector can be used to prepare a subunit vaccine with
CC other P. haemolytica antigens, e.g. fimbrial protein, plasmid
CC receptor or 50K outer membrane protein. The vaccines can be used
CC to protect cattle from respiratory diseases such as pneumonia, esp.
CC shipping fever pneumonia.
XX
CC See also R14481, 83,84 and 85.
XX
SQ Sequence 926 AA;

```

```

Query Match 59.2%; Score 2126; DB 12; Length 926;
Best Local Similarity 52.6%; Pred. No. 9.7e-135;
Matches 463; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

```

```

QY 111 SFEPTGAKKILLYIPONYOYDEQNGLODLVKAABELIEVOREERNINATQAOTSLGTT 170
Db 8 sfptgakkillyipnygydteqngldlvkaaeelgievgreernmlataqtslgtl 67
QY 171 QTAIGLTGRTIVSAPOIDKLQKTRAGQALGSASEIVQNAKAKVYVLSQISILGSVLA 230
Db 68 qtaigtltgrtivrivsapqidlqktkagqalgsaesivqnaakakvlylsgisilgsvla 127
QY 231 GMDLDEALONNSNOHALAKAGLELTSLJENINANSYKTLDEFEQOISOPGSKQNTKGLG 290
Db 128 gmdldeaiqnnsnqhalakagletnsljenlansvkldefeqsgfysklnkylg 187
QY 291 TLGDKLKNIGLDKAGLGLDVISGLSGATPAALVLADKNASTKRVGAGFELANQVYVNT 350
Db 188 tlgdtklnlgldkaglgldvlsqllsgataalvldknaastakvvgagfelanqvyvnt 247
QY 351 TKAVSSYIIAQRVAAGLSSTGPVAALIASTVSLAISPLAFAGIADKFNNAKSLESAERF 410
Db 248 tkavssyllagrvaaglsstgpvaalaaastvslaispladkfnhakslesyaerf 307
QY 411 KRLGYGDNILAEYQRTGTIDASVTAINATALAINGVSAAAA----- 454
Db 308 krlgygdnllaeayrgrgtldasvtaintaalaagvsaaagvsiasplallvsglt 367
QY 455 ----- 454
Db 368 gvisltllgyskgamfehvanklnkiveweknmhknfyfengydarylanldnmkflin 427
QY 455 ----- 454
Db 428 lnkelgaerviaitqgwdnigldaglsrlgkvisgkayvdafeegkhikadkivld 487

```

```

OY 455 ----- 454
Db 488 sangliidvnsqkaktqhl1ftfp1ltpgtehrervqtkyeytklnlnrvdswk1tdg 547
OY 455 ----- 454
Db 548 aasstfdlnvvgriq1eldnagnvktkethk1iaklggednvnfvsgstleidggyd 607
OY 455 ----- 454
Db 608 rvhysrgnygal1idatketeggyfvmrfveclgk1hvevsthtalvgnrreek1eyrbs 667
OY 455 ----- 454
Db 668 mnghagyytktlkaveel1gtsnhd1fkyskfndafnggdvd1dngdnrl1fgqk 727
OY 455 ----- 454
Db 728 gdd1ldgngdd1dgqkgndllhggkgdd1fvhrkqgdnd1ltdsgndk1sfsdn1k 787
OY 455 ----- 454
Db 788 d1tfekvkhnlv1tnskkekvt1qnfreadfakvpyk1atkdex1ee1lqngger1ts 847
OY 515 KOVDLIAKNGKITODELSKVVDNVELLKHKNVTNSLDKLISSVSATSSMDSRNVLV 574
Db 848 kvydd1iakngk1tdgelskvvdnyell1khsknvnsldk1lssvatssndsrnvl 907
OY 575 APTSMUDGSLSSLOFARGS 593
Db 908 aptsmidgslss1qfarg 926

```

RESULT 13

```

R34545 R34545 standard; Protein: 926 AA.
AC R34545:
DT 23-AUG-1993 (first entry)
DE Leukotoxin 352 produced from pAA352.
KW Vector: LKT 352; flanking; recombinant; antigen; somatostatin;
KW gonadotropin releasing hormone; rotavirus viral protein 4;
KW carrier protein; lactation; reproduction.
OS Pasteurella haemolytica.
PN MO9308290-A.
PD 29-APR-1993.
PE 15-OCT-1992; 92MO-CA00449.
PR 16-OCT-1991; 91US-0779171.
PR 14-OCT-1992; 92US-0960932.
PA (UYSA-) UNIV SASKATCHEWAN.
PI Hughes HPA, Potter AA, Redmond MJ;
DR N-PSDB: Q41317.
XX WPI: 1993-152482/18.
XX Immunological carrier system with enhanced immunogenicity -
XX comprises chimeric protein comprising leucotoxin peptide or
XX homologous protein fused to antigen esp. somatostatin or
XX gonadotropin releasing hormone
XX Disclosure; Fig 3; 95pp; English.
XX

```

CC Gene libraries of *P. haemolytica* A1 (strain B122) were constructed
 CC in lambda gtl1 and pUC13. Resulting clones were used to transform *E.*
 CC coli and individual colonies were pooled and screened for reaction
 CC with serum from a calf which had survived a *P. haemolytica* infection
 CC and that had been boosted with a conc. culture supernatant of *P.*
 CC *haemolytica* to increase anti-leukotoxin antibody levels. Positive
 CC colonies were screened for their ability to produce leukotoxin by
 CC incubating cell lysates with bovine neutrophils and measuring the
 CC release of lactate dehydrogenase from the neutrophils. A 4kb
 CC fragment was obtd. Progressively larger clones were isolated by
 CC chromosome walking to isolate full length recombinants of ca. 8kb,
 CC in pAA14. The clone was subjected to restriction enzyme digestion
 CC to yield two clones, one expressing truncated leukotoxin peptide at
 CC high levels and the other expressing the full length leukotoxin at
 CC low levels. The 3' end of the lktA gene from the full length clone
 CC was ligated to the truncated gene clone to yield plasmid pAA352. The
 CC clone was used to produce chimeric proteins by gene fusion with an
 CC antigen coding sequence, e.g. the coding sequence of somatostatin,
 CC gonadotropin releasing hormone or rotavirus viral protein 4, i.e.
 CC leukotoxin works as a carrier protein to bring about a larger
 CC immune response than the antigen alone. Immunisation with these
 CC antigens can regulate growth rate, lactation and reproductive efficiency.
 CC See also R34546-8.
 CC
 SO Sequence 926 AA:

Query Match 59.2%; Score 2126; DB 14; Length 926; 1
 Best Local Similarity 52.6%; Pred. No. 9,7e-135;
 Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

```

OY 111 SPFKTAKKIIILYIPONYOYDFEQNGLDLVKAAELG1EVOERRNNIATQTSLGIT 170
Db 8 sfpktagk11lyipnyqydeqng1qdlv1aaee1g1evgreern1ataqts1g1t 67
OY 171 QTAIGLTERGIYVSAPOIKLLOKTRKAGALGSAESIYONANKAKVLVLSG1S1LGSVLA 230
Db 68 qtaig1tergiy1sapq1k1l1qktaag1a1gsaes1vqan1kakv1s1g1s1g1sv1a 127
OY 231 GMDLDEALONNSQHALAKAGLELTNSLIENIANSVKTLDERGEQ1SOPGSKLQNTK1G 290
Db 128 gmd1dealg1nsq1halakag1elt1ns1lien1ansv1kt1der1geq1s1q1s1g1sk1gn1k1g 187
OY 291 TLGDKRKNTIGGDKAGLG1DV1SG1LSGAT1ALV1ADKNA1STAKKYGACFELANOVVNI 350
Db 188 tlgdk1kn1g1dkag1lg1dv1sg1lsg1ata1vlad1kna1stak1kvgag1felan1gvn1 247
OY 351 TKAVSSYILAO RVAAGLSTGPAALIASTVLA1SPLFAG1ADKFNHAK1S1ESYAEER 410
Db 248 tkavssy1laqv1aa1gst1cpvaal1astv1a1s1plaf1ag1adk1fnhaks1esyae1r 307
OY 411 KRLGYDGNMLAEYORG1GT1DASVTA1NTAL1AGVSA1AA----- 454
Db 308 krlgydgnd1laeyrg1gt1dasv1a1ntal1aag1vsa1aaag1v1asp1a1llvsg1t 367
OY 455 ----- 454
Db 368 gy1st1l1gys1kgam1feh1vank1h1nk1ivewek1nhg1kfyng1dary1an1lgdm1k1fl1n 427
OY 455 ----- 454
Db 428 l1kel1gaerv1a1ltq1qgd1n1lg1ag1sr1lge1k1v1sg1kay1vda1feeg1h1k1adk1lv1qd 487
OY 455 ----- 454
Db 488 sangliidvnsqkaktqhl1ftfp1ltpgtehrervqtkyeytklnlnrvdswk1tdg 547
OY 455 ----- 454
Db 548 aasstfdlnvvgriq1eldnagnvktkethk1iaklggednvnfvsgstleidggyd 607
OY 455 ----- 454

```

Db 608 rvhysrgnygalldatkebegsgyfvnrfvetgkalhevtsthalvgnreekleyrhs 667
 QY 455 ----- 454
 Db 668 mngnhagyytkdtlkaveeigtshndifkyskfnda fnggdvdtldgndgndrlfgyk 727
 QY 455 ----- 454
 Db 728 gddllldgngddfidggkgndllhggkgddlfvhrkggndlltdsgndklfsdsnlk 787
 QY 455 DLTFEYKXNHLVTTNSKKEKVTIQNMFREADPKKVPNKATDEKIEELTIGNGERITS 514
 Db 788 dlfevkvkhlvltnskkekvltqnmfreadfakvpykatdekfeelligngerits 847
 QY 515 KOYDDLIKNGKITODELSKVVDNVELKHSKNVTNSLDKLSSVSASFSSNDSRNLV 574
 Db 848 kyvddliakngkitqdeliskvvdnyellkhsknvnsldklssvsafssndsrnlv 907
 QY 575 APTSMLDOSLSSLOFARGS 593
 Db 908 aptsmldgsslsqfargs 926
 RESULT 14
 R50291
 ID R50291 standard; Protein: 926 AA.
 AC R50291;
 DT 06-OCT-1994 (first entry)
 DE Recombinant leukotoxin from plasmid PAA352.
 KM Vaccine: outer membrane protein; OMP: Haemophilus somnus;
 KM Iron regulated protein; leukotoxin; Pasteurella haemolytica;
 KM LKT352.
 OS Pasteurella haemolytica A1 (strain B122).
 PN CA2099707-A.
 PD 03-JAN-1994.
 PF 29-JUN-1993; 93CA-2099707.
 PR 02-JUL-1992; 92US-0908253.
 PA (UYSA-) UNITV SASKATCHEWAN.
 PI Harland RJ, Potter AA;
 DR MPI: 1994-092909/12.
 DR N-PSDB; Q44760.
 XX Haemophilus somnus outer membrane protein extract -
 PT enriched with iron-regulated proteins, opt. contg.
 PT Leuco:toxin antigens, for use as vaccine
 XX
 PS Claim 5: Fig 5: 78pp; English.
 CC A vaccine comprising an outer membrane protein (OMP) extract of
 CC Haemophilus somnus enriched with iron regulated proteins is new.
 CC The vaccine pref. further comprises an immunogenic leukotoxin
 CC polypeptide, esp. an immunogenic pasteurella haemolytica leukotoxin
 CC homologous to LKT352. Example 1.2 describes the prodn. of
 CC P. haemolytica recombinant leukotoxin from PAA352.
 CC Two expression constructs were made. One, PAA342, contained the
 CC 5'-AhaiI fragment of the ltkA gene, while the other, PAA345,
 CC contained the entire ltkA gene. PAA342 expressed a truncated
 CC leukotoxin peptide at high levels, while PAA345 expressed full
 CC length leukotoxin at very low levels. Therefore, the 3' end
 CC of the ltkA gene was ligated into PAA342, yielding plasmid PAA352.
 CC LKT352 or new leukotoxin is 98% homologous to authentic

CC leukotoxin.
 CC NB: the protein sequence in Fig 5 comprises 926 amino acids,
 CC however this protein is described in the text as having
 CC 931 amino acids.
 XX
 SQ Sequence 926 AA;
 Query Match 59.2%; Score 2126; DB 15; Length 926;
 Best Local Similarity 52.6%; Pred. No. 9.7e-135;
 Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;
 QY 111 SFPKTKAKKILLYIPONYOYDTEQNGLODLYKAEELGIEVOREERNNTATQTSIGTI 170
 Db 8 sfktgkakkililypnygydteqnglqdlvkaaeelgvevgreenmlatqtsigt 67
 QY 171 QTAIGLTERIVASQIDKLLOKTKAGQALSGAESIVQANAKKIVLSIGTISVLA 1230
 Db 68 qtaiglttergvisapqidlqktkaggalgsaesivqnaankakvlsigislvla 127
 QY 231 GMDLDEALONSNQHLAKAGLELTNSLIENIANSVKTTDEFEEOISOFCKLQNIKGLG 290
 Db 128 gmdldealqnsnqhalakagletnslleiansvkltdeifeqisqfsgkignlkgy 187
 QY 291 TLGDKLKNIGGLDKAGLGLDVISGLSGATPAALVLADKNASTAKKYGAGFELANQVYGI 350
 Db 188 tlgdklknigldkaglgldvisglsgatpaalvladknaastakkyvga fela ngyvgnl 247
 QY 351 TKAVSSYIIAQRYAAGLSSGPYAAALIASTVSLAISPLAFAGIADKFNNAKSLSEAEER 410
 Db 248 tkavssyllaqryaaaglstgpyaaalastvslaisplafagiadkfnhakslesyaerf 307
 QY 411 KRLGYGDNLLAEYRGSTGTIDASYTAINTALAIAGVSAANA----- 454
 Db 308 krlgygdnllaeqrgtltidasvtaintalaiaigvsaanaagvsia spiallvsglt 367
 QY 455 ----- 454
 Db 368 gvalstllqyskgamfehvanklnhkliveweknhgknfyfengydarylamlqdnmkflln 427
 QY 455 ----- 454
 Db 428 lnkelaervialtqgqwdnigdlaglsrlgekvlsqkayvdafeegkhikadklvqlid 487
 QY 455 ----- 454
 Db 488 sangliavnsqakqghllfrpplltpgtehrerqvtkyeyitklnhrvdsakitdg 547
 QY 455 ----- 454
 Db 548 aasftdltnvgigldelnagnvcktketkilaiklgegdhvfvgsgteldggygd 607
 QY 455 ----- 454
 Db 608 rvhysrgnygalldatkebegsgyfvnrfvetgkalhevtsthalvgnreekleyrhs 667
 QY 455 ----- 454
 Db 668 mngnhagyytkdtlkaveeigtshndifkyskfnda fnggdvdtldgndgndrlfgyk 727
 QY 455 ----- 454
 Db 728 gddllldgngddfidggkgndllhggkgddlfvhrkggndlltdsgndklfsdsnlk 787
 QY 455 DLTFEYKXNHLVTTNSKKEKVTIQNMFREADPKKVPNKATDEKIEELTIGNGERITS 514
 Db 788 dlfevkvkhlvltnskkekvltqnmfreadfakvpykatdekfeelligngerits 847
 QY 515 KOYDDLIKNGKITODELSKVVDNVELKHSKNVTNSLDKLSSVSASFSSNDSRNLV 574
 Db 848 kyvddliakngkitqdeliskvvdnyellkhsknvnsldklssvsafssndsrnlv 907
 QY 575 APTSMLDOSLSSLOFARGS 593

```

Db      908 aptsmldgslslqfargs 926
|||||
RESULT 15
W03945
ID      W03945 standard; Protein; 926 AA.
XX
AC      W03945;
XX
DT      20-NOV-1996 (first entry)
XX
DE      P. haemolytica truncated leukotoxin (LKT352).
XX
KW      Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
KW      fusion protein; immunogen; vaccine; fertility control;
KW      contraceptive; sterilisation; plasmid pAA352.
XX
OS      Pasteurella haemolytica A1 strain B122.
XX
PN      W09624675-A1.
XX
PD      15-AUG-1996.
XX
PF      24-JAN-1996; 96MO-CA00049.
XX
PR      10-FEB-1995; 95US-0387156.
XX
PA      (UYSA-) UNIV SASKATCHEWAN.
XX
PI      Manns JG, Potter AA;
XX
DR      WPI: 1996-384447/38.
XX      N-PSDB; T37179.
XX
PT      Gonadotropin-releasing hormone multimer fusion proteins - with
PT      leukotoxin polypeptide for increased immunogenicity, useful in
PT      antifertility vaccine prodn.
XX
PS      Example 1; Fig 3A-3I; 87pp; English.
XX
CC      A truncated leukotoxin (W03945), LKT 352, lacks the cytotoxic
CC      portion of the native protein from Pasteurella haemolytica. It is
CC      the product of plasmid pAA352 which carries a truncated lktA gene
CC      (T37179). A fusion protein (W03942) between LKT352 and a
CC      gonadotropin releasing hormone tetramer can be expressed in
CC      Escherichia coli. This is useful as a vaccine for fertility
CC      control, partic. immunological sterilisation of domestic or
CC      farm animals.
XX
SQ      Sequence 926 AA:
Query Match 59.2%; Score 2126; DB 17; Length 926;
Best Local Similarity 52.6%; Pred. No. 9.7e-135;
Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;
OY      111 SPFKTGAKKIIYIPONTQYDTREGNGIADLYKAABELGIEVOREBRNNIATPQTSIGT 170
DB      8 sfpktygakkililypqnyqdydegngldlvkaaeelgievgreenmlataqtslgtl 67
OY      171 OTAIGTGERGIYLSAPOIDKLOKTKAGALGASISIVONANKAKTVLSGIOSISVLA 230
DB      68 qtaigttergiyvisapqdkllqtkcagaglsaeslvqnankaktvlsqslgsvla 127
OY      231 GMDLDEALONNSQHALAKAGLELTNSLTENIANSVKTLDEFGEQISQFGSKLQNIKGLG 290
DB      128 gmdldealgqnsnqhalakagletnslieniensvktldefgeqisqfsgsklqnkglg 187
OY      291 TLGDKLKNIGGIDKAGLGIDVISGLSGTALVLAADKNAKTAKKVAGAGFELANOVGNI 350
DB      188 tlgdklknigldkaglgidvisgllsgataalvladknaastakkvagafelanvvgni 247

```

```

OY      351 TKAVSSYIIAORVAAGLSTSTGPVALIASTVSLAISPLAFAGIADKFNHAKLSYAEERF 410
DB      248 tkavssyilaqrvaagllststgpvaaliastvsalaspiafagidackfmhakslesy 307
OY      411 KRLGYDGNMLAEYORGTGTIDASVTAINTAALAIAGVSAAA----- 454
DB      308 krlgydgnmlaeygrgtictidasvtaintalaiagiysaaagsvlaspiallvsgit 367
OY      455 ----- 454
DB      368 gyistllgyskqamfehvaanklhnkiveweknhgknyfengydarylanlqdmkflin 427
OY      455 ----- 454
DB      428 lnkelaervialetqgwdmngldaglsrlgekvlsqkayvdafeegkhkacklvql 487
OY      455 ----- 454
DB      488 saanglidsnsgkaktqhlfrtplltpgtetrvrvqgkyeyicklnrvdsklidg 547
OY      455 ----- 454
DB      548 aasstfdlnuvvgigieladnagnvltkethkilaigeddnvfvsgteldggyd 607
OY      455 ----- 454
DB      608 rvhysrgnygaltdaketegegsytlvrfvetgkalhevsthalvgnreekleyrhs 667
OY      455 ----- 454
DB      668 nqghagytkdtkaveqilgstshndlfksgkfndafngdgvdtldgndndrlfpgk 727
OY      455 ----- 454
DB      728 gddlldgngddfidgqkgnldlhggkgddlfvrhkgdgnldltdsgndklsfsgn 787
OY      455 DLTPFKVKHNLVITTSKKKEKVTIOMWPREADFAKVPYKATKDKIEEITIGONGERTS 514
DB      788 dltpfkvhnlvlttskkkvtiomwpreadfakevpykakdekleeilgqngerlts 847
OY      515 KOVDLIRKNGKITQDELKVVVDNYELIKHSKNTNSLDKLISVSASFSTSNDSRNVLV 574
DB      848 kqvddlirkngkitqdelskvvdnyellhsknvtnsldklisvsafstsnv 907
OY      575 APTSMUDQSLSLQFARGS 593
DB      908 aptsmldgslslqfargs 926

```

Search completed: March 2, 2001, 10:53:59
Job time: 924 sec

THIS PAGE BLANK (USPTO)


```
Db 1 MATVIDRSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQ 60
QY 61 HWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSFPKTKAKKI 120
Db 61 HWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSFPKTKAKKI 120
QY 121 ILYIPONYOYDTBOGNGLDOLVKAABELGIEVOREERNINATATOTSIGTQTATIGLTERG 180
Db 121 ILYIPONYOYDTBOGNGLDOLVKAABELGIEVOREERNINATATOTSIGTQTATIGLTERG 180
QY 181 IYLSAPOIDKLOKTRAGQALSAESIYONANAKTYLSGTSIGSLVLAGMDLDELALON 240
Db 181 IYLSAPOIDKLOKTRAGQALSAESIYONANAKTYLSGTSIGSLVLAGMDLDELALON 240
QY 241 NSNOHALAKAGLELJNSLIENIANSVKTTIDFEQEIOISQFSGSKLONIKGLTGLDKLNIG 300
Db 241 NSNOHALAKAGLELJNSLIENIANSVKTTIDFEQEIOISQFSGSKLONIKGLTGLDKLNIG 300
QY 301 GLDKAGIGLDVIGSLGATPAALVLDKNASTAKKVGAGFELANOVVGNITTKAVSSYILA 360
Db 301 GLDKAGIGLDVIGSLGATPAALVLDKNASTAKKVGAGFELANOVVGNITTKAVSSYILA 360
QY 361 QRYAAGLSSTGPAALIASVSLAISPLAFAGIADKFENHAKSLESYAERPKKLGIOGDNL 420
Db 361 QRYAAGLSSTGPAALIASVSLAISPLAFAGIADKFENHAKSLESYAERPKKLGIOGDNL 420
QY 421 LAEYORGTGTIDASVTAINTALAAIAGVSAAADLTFEKKVKNHLYITNSKKKEKVTIOMN 480
Db 421 LAEYORGTGTIDASVTAINTALAAIAGVSAAADLTFEKKVKNHLYITNSKKKEKVTIOMN 480
QY 481 FRADFAKEVPNTKATKDEKIEEIIQNGERTSKOYDLDLIAKNGKITODELSKYVDNY 540
Db 481 FRADFAKEVPNTKATKDEKIEEIIQNGERTSKOYDLDLIAKNGKITODELSKYVDNY 540
QY 541 ELKHKSNTVNSLDKLISVSATSSNDSRNVLVAPTSMIDSLSSLOFARQSQHWSYGL 600
Db 541 ELKHKSNTVNSLDKLISVSATSSNDSRNVLVAPTSMIDSLSSLOFARQSQHWSYGL 600
QY 601 RPSGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMS 660
Db 601 RPSGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMS 660
QY 661 SYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSS 695
Db 661 SYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSS 695

RESULT 2
US-09-306-689-13
; Sequence 13, Application US/09306689B
; GENERAL INFORMATION:
; APPLICANT: Robbins, Sarah C.
; TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
; TITLE OF INVENTION: ANIMALS
; FILE REFERENCE: 9001-0047
; CURRENT APPLICATION NUMBER: US/09/306,689B
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: US 60/088,024
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 13
; LENGTH: 695
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; OTHER INFORMATION: defines a leukotoxin GnRH chimera
US-09-306-689-13

Query Match 100.0%; Score 3591; DB 17; Length 695;
Best Local Similarity 100.0%; Pred. No. 3,le-281;
```

```
Matches 695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATVIDRSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQ 60
Db 1 MATVIDRSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQ 60
QY 61 HWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSFPKTKAKKI 120
Db 61 HWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSFPKTKAKKI 120
QY 121 ILYIPONYOYDTBOGNGLDOLVKAABELGIEVOREERNINATATOTSIGTQTATIGLTERG 180
Db 121 ILYIPONYOYDTBOGNGLDOLVKAABELGIEVOREERNINATATOTSIGTQTATIGLTERG 180
QY 181 IYLSAPOIDKLOKTRAGQALSAESIYONANAKTYLSGTSIGSLVLAGMDLDELALON 240
Db 181 IYLSAPOIDKLOKTRAGQALSAESIYONANAKTYLSGTSIGSLVLAGMDLDELALON 240
QY 241 NSNOHALAKAGLELJNSLIENIANSVKTTIDFEQEIOISQFSGSKLONIKGLTGLDKLNIG 300
Db 241 NSNOHALAKAGLELJNSLIENIANSVKTTIDFEQEIOISQFSGSKLONIKGLTGLDKLNIG 300
QY 301 GLDKAGIGLDVIGSLGATPAALVLDKNASTAKKVGAGFELANOVVGNITTKAVSSYILA 360
Db 301 GLDKAGIGLDVIGSLGATPAALVLDKNASTAKKVGAGFELANOVVGNITTKAVSSYILA 360
QY 361 QRYAAGLSSTGPAALIASVSLAISPLAFAGIADKFENHAKSLESYAERPKKLGIOGDNL 420
Db 361 QRYAAGLSSTGPAALIASVSLAISPLAFAGIADKFENHAKSLESYAERPKKLGIOGDNL 420
QY 421 LAEYORGTGTIDASVTAINTALAAIAGVSAAADLTFEKKVKNHLYITNSKKKEKVTIOMN 480
Db 421 LAEYORGTGTIDASVTAINTALAAIAGVSAAADLTFEKKVKNHLYITNSKKKEKVTIOMN 480
QY 481 FRADFAKEVPNTKATKDEKIEEIIQNGERTSKOYDLDLIAKNGKITODELSKYVDNY 540
Db 481 FRADFAKEVPNTKATKDEKIEEIIQNGERTSKOYDLDLIAKNGKITODELSKYVDNY 540
QY 541 ELKHKSNTVNSLDKLISVSATSSNDSRNVLVAPTSMIDSLSSLOFARQSQHWSYGL 600
Db 541 ELKHKSNTVNSLDKLISVSATSSNDSRNVLVAPTSMIDSLSSLOFARQSQHWSYGL 600
QY 601 RPSGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMS 660
Db 601 RPSGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMS 660
QY 661 SYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSS 695
Db 661 SYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSS 695

RESULT 3
US-09-383-912-16
; Sequence 16, Application US/09383912
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANNS, JOHN G.
; TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
```



```

: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/694,865
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: MCCracken, THOMAS P.
: REGISTRATION NUMBER: 38,548
: REFERENCE/DOCKET NUMBER: 9001-0016.22
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415)327-3400
: TELEFAX: (415)327-3231
: INFORMATION FOR SEQ ID NO: 16:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 699 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-09-383-912-16

```

```

Query Match      100.0%; Score 3591; DB 17; Length 699;
Best Local Similarity 100.0%; Pred. No. 3.1e-281;
Matches 695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 MATVIDRSQHMSYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQ
DB 1 MATVIDRSQHMSYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQ
OY 61 HMSYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQ
DB 61 HMSYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQ
OY 121 ILIYPONTQYDTBOGNGIQLDYLKAAEELGIVOREERNNTAQTSLGTITGTLTBERG
DB 121 ILIYPONTQYDTBOGNGIQLDYLKAAEELGIVOREERNNTAQTSLGTITGTLTBERG
OY 181 IYLSAPQIDKLQTKKAGALGSAESIYQNNKAKTVLSGIOSILGSLVLAQMDDEALQN
DB 181 IYLSAPQIDKLQTKKAGALGSAESIYQNNKAKTVLSGIOSILGSLVLAQMDDEALQN
OY 241 NSNQHAKAGLELTLNLIENIANSVKTLDFEGEIOISQFSKLNKIGLGLTGLKKNIG
DB 241 NSNQHAKAGLELTLNLIENIANSVKTLDFEGEIOISQFSKLNKIGLGLTGLKKNIG
OY 301 GLDKAGLGLDVISGLSGATATLADKNASTAKKVGAGFELANOVGNITKAVSSYTLA
DB 301 GLDKAGLGLDVISGLSGATATLADKNASTAKKVGAGFELANOVGNITKAVSSYTLA
OY 361 ORVAAGLSSTGPVAALIASTVSLAISPFAAGIADKFNHAKSLESYARFKKLGIDGNL
DB 361 ORVAAGLSSTGPVAALIASTVSLAISPFAAGIADKFNHAKSLESYARFKKLGIDGNL
OY 421 LAEYORGTGTIDASTATINTALAAIAGVSAAADLTFEKYNHNLVITNSKKEKTYIONW
DB 421 LAEYORGTGTIDASTATINTALAAIAGVSAAADLTFEKYNHNLVITNSKKEKTYIONW
OY 481 FREDFAKEVNYKATKDEKIEEIIIGNGERITSKQVDDLIAGKNGKITODELSKYVDNY
DB 481 FREDFAKEVNYKATKDEKIEEIIIGNGERITSKQVDDLIAGKNGKITODELSKYVDNY
OY 541 ELHKRSKNTVSLDKLSSVSAFTSSNDSRNVLVAPTSMLDOSLSSLOFARSGSHMSYGL
DB 541 ELHKRSKNTVSLDKLSSVSAFTSSNDSRNVLVAPTSMLDOSLSSLOFARSGSHMSYGL
OY 601 RPSGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDM
DB 601 RPSGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDM
OY 661 SYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQ
DB 661 SYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQ

```

```

RESULT 4
US-09-383-912-10
: Sequence 10, Application US/09383912
: GENERAL INFORMATION:
: APPLICANT: POTTER, ANDREW A.
: APPLICANT: MANN, JOHN G.
: TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: REED & ROBINS LLP
: STREET: 285 HAMILTON AVENUE, SUITE 200
: CITY: PALO ALTO
: STATE: CA
: COUNTRY: USA
: ZIP: 94301
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/383,912
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/694,865
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: MCCracken, THOMAS P.
: REGISTRATION NUMBER: 38,548
: REFERENCE/DOCKET NUMBER: 9001-0016.22
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415)327-3400
: TELEFAX: (415)327-3231
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 544 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-09-383-912-10

```

```

Query Match      73.6%; Score 2642.5; DB 17; Length 544;
Best Local Similarity 99.4%; Pred. No. 9.4e-205;
Matches 534; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```

```

OY 111 SFPKGAKKILTYIPONTQYDTBOGNGIQLDYLKAAEELGIVOREERNNTAQTSLGTI
DB 8 SFPKGAKKILTYIPONTQYDTBOGNGIQLDYLKAAEELGIVOREERNNTAQTSLGTI
OY 171 QTAIGLTERGIVLSAPQIDKLQTKKAGALGSAESIYQNNKAKTVLSGIOSILGSLVLA
DB 68 QTAIGLTERGIVLSAPQIDKLQTKKAGALGSAESIYQNNKAKTVLSGIOSILGSLVLA
OY 231 GMDLDEALONNSNQHAKAGLELTLNLIENIANSVKTLDFEGEIOISQFSKLNKIGL
DB 128 GMDLDEALONNSNQHAKAGLELTLNLIENIANSVKTLDFEGEIOISQFSKLNKIGL
OY 291 TLSPDKLNKIGLDAAGLGLDVISGLSGATATLADKNASTAKKVGAGFELANOVGNIT
DB 188 TLSPDKLNKIGLDAAGLGLDVISGLSGATATLADKNASTAKKVGAGFELANOVGNIT
OY 351 TKAVSSYTLAORVAAGLSSTGPVAALIASTVSLAISPFAAGIADKFNHAKSLESYARF
DB 248 TKAVSSYTLAORVAAGLSSTGPVAALIASTVSLAISPFAAGIADKFNHAKSLESYARF
OY 411 KKLGDGNLLAEYORGTGTIDASTATINTALAAIAGVSAAAA---DLTFEKYHNLVI
DB 308 KKLGDGNLLAEYORGTGTIDASTATINTALAAIAGVSAAAA---DLTFEKYHNLVI
OY 468 TNSKKEKTYIONWTFADFAKEVNYKATKDEKIEEIIIGNGERITSKQVDDLIAGKNGK
DB 468 TNSKKEKTYIONWTFADFAKEVNYKATKDEKIEEIIIGNGERITSKQVDDLIAGKNGK

```

```

Db 368 TNSKKEKVIQNNFREADFAKEVPNKATKDEKIEIEIIGONGERITSQYDDLIANGNGK 427
QY 528 ITODELSKYVDNELLKHSKNVTNSLDKLSSVSAPFSSNDNRNVLPAPSMIDQSLSSL 567
Db 428 ITODELSKYVDNELLKHSKNVTNSLDKLSSVSAPFSSNDNRNVLPAPSMIDQSLSSL 487
QY 588 OFARQSQHMSYGLRPGSSQDMSYGLRPGSSQHMSYGLRPGSSQDMSYGLRPGGS 644
Db 488 OFARQSQHMSYGLRPGSSQDMSYGLRPGSSQHMSYGLRPGSSQDMSYGLRPGGS 544

RESULT 5
US-09-383-912-8
; Sequence 8, Application US/09383912
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 977 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-383-912-8

Query Match 67.6%; Score 2426; DB 17; Length 977;
Best Local Similarity 55.1%; Pred. No. 7,8e-187;
Matches 534; Conservative 0; Mismatches 0; Indels 436; Gaps 1;
QY 111 SPKTKAKKIIILYIPONYQYDTEQNGLODLVKAEBLGIEVOREERNNTATQTSIGTI 170
Db 8 SPKTKAKKIIILYIPONYQYDTEQNGLODLVKAEBLGIEVOREERNNTATQTSIGTI 67
QY 171 QTVIGLTERGIVASAPQIDLLQKTAGAGALGSAESIYONANKAKTVLGSIGISLVLA 230
Db 68 QTVIGLTERGIVASAPQIDLLQKTAGAGALGSAESIYONANKAKTVLGSIGISLVLA 127
QY 231 GMDLDEALQNSNQHALAKAGLELTNSLIENINSYKTLDEPGEQISQFSGSKQNTKGLG 290
Db 128 GMDLDEALQNSNQHALAKAGLELTNSLIENINSYKTLDEPGEQISQFSGSKQNTKGLG 187
QY 291 TIGDKLKNTIGGLDKAGIGLDIVISGLLSGATPAALVLADKNASTAKKVGAGFELANQVYVNT 350

```

```

Db 188 TIGDKLKNTIGGLDKAGIGLDIVISGLLSGATPAALVLADKNASTAKKVGAGFELANQVYVNT 247
QY 351 TKAVSSYTLIAQRYAAGISSTGPAALIASTVSLATISPLAFAGIADKFNNAKSLSEYAE 410
Db 248 TKAVSSYTLIAQRYAAGISSTGPAALIASTVSLATISPLAFAGIADKFNNAKSLSEYAE 307
QY 411 KKLGDGDNLLAEYORGTGTIDASVTAINPTALAIAGVSAANA----- 454
Db 308 KKLGDGDNLLAEYORGTGTIDASVTAINPTALAIAGVSAANAAGSVIASPIALLVSGIT 367
QY 455 ----- 454
Db 368 GVISTILLQYSQAMFEHVANKIHKIVEMKNNHKNPENGYDARYLANLDNMKFLIN 427
QY 455 ----- 454
Db 428 LNKELDAERYIALITQOQWNNIGDLAGISRLGKVLGSKAYVDAFEGKHIRAKDLVOLD 487
QY 455 ----- 454
Db 488 SANGIIDVNSGKAKTQHILFTPTLLPGTEHREYOTGKEYITKLNINRVDWMKITDG 547
QY 455 ----- 454
Db 548 AASTFDLTNVVORIGIELDNAGNVTKETKLIARKLEGDDNVFVSGTTEIDGEGYD 607
QY 455 ----- 454
Db 608 RVHYSKNGALITIDATKETEQSYTVNREVEFGKALHEVTSHTHTALVGNREKIEYRHS 667
QY 455 ----- 454
Db 668 NNGHAGYYTKDTLKAVEELIGTSHNDIFPKSGKFNDAFNGDGVDTIDGNDRLFGGK 727
QY 455 ----- 454
Db 728 GDDILDGNGDDFIDGKGNLDLHGKGDDIFVHRKGDNDIITDSGDNKLSFSDNKK 787
QY 455 DLTFEYKHNLTNTSKKKEKVTIONMFREADFAKEVPNKATKDEKIEIEIIGONGERITS 514
Db 788 DLTFEYKHNLTNTSKKKEKVTIONMFREADFAKEVPNKATKDEKIEIEIIGONGERITS 847
QY 515 KQYDDLIANGNGKITODELSKYVDNELLKHSKNVTNSLDKLSSVSAPFSSNDNRNVLP 574
Db 848 KQYDDLIANGNGKITODELSKYVDNELLKHSKNVTNSLDKLSSVSAPFSSNDNRNVLP 907
QY 575 APTSMIDQSLSSIQFARQSQHMSYGLRPGSSQDMSYGLRPGSSQHMSYGLRPGGSQD 634
Db 908 APTSMIDQSLSSIQFARQSQHMSYGLRPGSSQDMSYGLRPGSSQHMSYGLRPGGSQD 967
QY 635 WSYGLRPGGS 644
Db 968 WSYGLRPGGS 977

RESULT 6
US-09-252-149A-26
; Sequence 26, Application US/09252149A
; GENERAL INFORMATION:
; APPLICANT: Barker, Christopher A.
; APPLICANT: Morsey, Mohamed
; TITLE OF INVENTION: IMMUNOLOGICAL METHODS TO MODULAR MYOSTATIN IN
; FILE REFERENCE: 9001-0042
; CURRENT APPLICATION NUMBER: US/09/252,149A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/075,213
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 26
; LENGTH: 490
; TYPE: PRT

```

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: leukotoxin polypeptide carrier
OTHER INFORMATION: Figures 15A-15D
US-09-252-149A-26

Query Match 65.6%; Score 2354; DB 16; Length 490;
Best Local Similarity 100.0%; Pred. No. 1.7e-181;
Matches 483; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 SPPKTKAKKIIILYIPQNTQYDTEGNGGLDLYKAAEELGIEVQREERNINATAQTSLSGTI 170
DB 8 SPPKTKAKKIIILYIPQNTQYDTEGNGGLDLYKAAEELGIEVQREERNINATAQTSLSGTI 67
QY 171 QTAIGLTERGIVLSAPQIDKLLQTKRAGALGASISIVONAKKATVLSGISISVLA 230
DB 68 QTAIGLTERGIVLSAPQIDKLLQTKRAGALGASISIVONAKKATVLSGISISVLA 127
QY 231 GMDLDEALQNNNSNOHALAKAGLELTNSLIENTANSVKTLDEGEQISQPSKLNKIGLG 290
DB 128 GMDLDEALQNNNSNOHALAKAGLELTNSLIENTANSVKTLDEGEQISQPSKLNKIGLG 187
QY 291 TLGDKLKNIGLDKAGLGIDVLSGLSGATATVLDKNASTAKKVGAGFELANQVGN 350
DB 188 TLGDKLKNIGLDKAGLGIDVLSGLSGATATVLDKNASTAKKVGAGFELANQVGN 247
QY 351 TKAVSSYLLAQRYAAGLSSTGPAALLASTVSLAISPLAFAGIADKFHNAKSLESYAERF 410
DB 248 TKAVSSYLLAQRYAAGLSSTGPAALLASTVSLAISPLAFAGIADKFHNAKSLESYAERF 307
QY 411 KRLGYDGNLLAEYORGCTIDASYTAINTALAAIAGVSAADLTPEKYKHNLYTNS 470
DB 308 KRLGYDGNLLAEYORGCTIDASYTAINTALAAIAGVSAADLTPEKYKHNLYTNS 367
QY 471 KKEKVTIONMFREAPFAFEVPMYKATKDEKIEIIGONGERTSKQVODLLAKNGKITQ 530
DB 368 KKEKVTIONMFREAPFAFEVPMYKATKDEKIEIIGONGERTSKQVODLLAKNGKITQ 427
QY 531 DELSKRVNDYELLKHSKNVTNSLDKLSSVSFTSSNDSRNVLVAPTSMLODLSLSLOFA 590
DB 428 DELSKRVNDYELLKHSKNVTNSLDKLSSVSFTSSNDSRNVLVAPTSMLODLSLSLOFA 487
QY 591 RGS 593
DB 488 RGS 490

RESULT 7

US-09-252-149B-26
Sequence 26, Application US/09252149B
GENERAL INFORMATION:
APPLICANT: Barker, Christopher A.
APPLICANT: Morsey, Mohamad
TITLE OF INVENTION: IMMUNOLOGICAL METHODS TO MODULAR MYOSTATIN IN
TITLE OF INVENTION: VERTEBRATE SUBJECTS
FILE REFERENCE: 9001-0042
CURRENT APPLICATION NUMBER: US/09/252,149B
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: 60/075,213
PRIOR FILING DATE: 1998-02-19
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 26
LENGTH: 490
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: leukotoxin polypeptide carrier
US-09-252-149B-26

Query Match 65.6%; Score 2354; DB 27; Length 490;
Best Local Similarity 100.0%; Pred. No. 1.7e-181;
Matches 483; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 SPPKTKAKKIIILYIPQNTQYDTEGNGGLDLYKAAEELGIEVQREERNINATAQTSLSGTI 170
DB 8 SPPKTKAKKIIILYIPQNTQYDTEGNGGLDLYKAAEELGIEVQREERNINATAQTSLSGTI 67
QY 171 QTAIGLTERGIVLSAPQIDKLLQTKRAGALGASISIVONAKKATVLSGISISVLA 230
DB 68 QTAIGLTERGIVLSAPQIDKLLQTKRAGALGASISIVONAKKATVLSGISISVLA 127
QY 231 GMDLDEALQNNNSNOHALAKAGLELTNSLIENTANSVKTLDEGEQISQPSKLNKIGLG 290
DB 128 GMDLDEALQNNNSNOHALAKAGLELTNSLIENTANSVKTLDEGEQISQPSKLNKIGLG 187
QY 291 TLGDKLKNIGLDKAGLGIDVLSGLSGATATVLDKNASTAKKVGAGFELANQVGN 350
DB 188 TLGDKLKNIGLDKAGLGIDVLSGLSGATATVLDKNASTAKKVGAGFELANQVGN 247
QY 351 TKAVSSYLLAQRYAAGLSSTGPAALLASTVSLAISPLAFAGIADKFHNAKSLESYAERF 410
DB 248 TKAVSSYLLAQRYAAGLSSTGPAALLASTVSLAISPLAFAGIADKFHNAKSLESYAERF 307
QY 411 KRLGYDGNLLAEYORGCTIDASYTAINTALAAIAGVSAADLTPEKYKHNLYTNS 470
DB 308 KRLGYDGNLLAEYORGCTIDASYTAINTALAAIAGVSAADLTPEKYKHNLYTNS 367
QY 471 KKEKVTIONMFREAPFAFEVPMYKATKDEKIEIIGONGERTSKQVODLLAKNGKITQ 530
DB 368 KKEKVTIONMFREAPFAFEVPMYKATKDEKIEIIGONGERTSKQVODLLAKNGKITQ 427
QY 531 DELSKRVNDYELLKHSKNVTNSLDKLSSVSFTSSNDSRNVLVAPTSMLODLSLSLOFA 590
DB 428 DELSKRVNDYELLKHSKNVTNSLDKLSSVSFTSSNDSRNVLVAPTSMLODLSLSLOFA 487
QY 591 RGS 593
DB 488 RGS 490

RESULT 8

US-07-779-171-10
Sequence 10, Application US/07779171
GENERAL INFORMATION:
APPLICANT: Potter, Andrew A.
APPLICANT: Redmond, Mark J.
APPLICANT: Hughes, Huw P. A.
TITLE OF INVENTION: ENHANCED IMMUNOCENTRICITY USING
TITLE OF INVENTION: PASTEURILLA HEAMOLYTICA LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 545 Middlefield Road, suite 200
CITY: Menlo Park
STATE: California
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/779,171
FILING DATE: 19911016
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Robins, Roberta L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 29310-2001600
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-327-7250


```

Db 248 TRAVSSYLLAQRVAAGLSTGPAALIASTVSLAISPLAFAGIAQENAKSLESYAERF
OY 411 KRLGVDGNLLAEYRGSTIDASTATATATATAIAGVSAASA-----
Db 308 KRLGVDGNLLAEYRGSTIDASTATATATAIAGVSAASAASVIAPIALLVSGIT
OY 455 -----
Db 368 GVIITLLQSKQAMFEHVANKIHKKIVEMKNNHGNKTFENGIDARYLANLQDNMKFLIN
OY 455 -----
Db 428 LNKELQAEERVAITQOQMNINIGLAGISRLGEKVLGSKAYDAFEQKHAKADLYOLD
OY 455 -----
Db 488 SANGIIDVNSNGAKTQHILFRTPLTPGTEHREVRQYTGKEYITKLNINRVDNWKITDG
OY 455 -----
Db 548 AASTFEDLTNVQRIEILDNAGNVTKTETKIIAKLGEGDNNVFGSGTTEIDGEGYD
OY 455 -----
Db 608 RVHYSRGNYGALTIDATKETEGSYTVNRFVETGKALHEVSTHTALVGNREKIEYRHS
OY 455 -----
Db 668 NNQHHAGYTTKDTLKAVEEIIIGTSHNDIFKSGFNDAFNGDGVDTIDGNDNRLFGGK
OY 455 -----
Db 728 GDDIILGNGDDFDIGGKGNLLHGKGGDIFVHRKGGNDIITDSGNDKLSFSDSNLK
OY 455 -----
Db 788 DLTFERKVNHLVITNSKKEKVTIQNMFRADPAKEVPYNTKATDEKIEIIGONGERITS
OY 455 -----
Db 847 DLTFERKVNHLVITNSKKEKVTIQNMFRADPAKEVPYNTKATDEKIEIIGONGERITS
OY 515 KOVDLLIAGKNGKITODELSKVVNDYELLKHSKNTNSIDKLISVSASFSSNDSRNVLY
Db 848 KOVDLLIAGKNGKITODELSKVVNDYELLKHSKNTNSIDKLISVSASFSSNDSRNVLY
OY 575 APTSMIDQSLSLQPARSGOHWSYGLRPG 603
Db 908 APTSMIDQSLSLQPARSGOHWSYGLRPG 936

```

```

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Robins, Roberta L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 29310-2001600
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-327-7250
TELEFAX: 415-327-2951
TELEX: 706141
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 926 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-779-171-14

Query Match 59.2%; Score 2126; DB 3; Length 926;
Best Local Similarity 52.6%; Pred. No. 1.3e-162;
Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

OY 111 SEPKTGARKIILYIPQNYQYDTEQNGHLDLVKAAEEIGIEVQREERNINATAQTSIGTI 170
Db 8 SEPKTGARKIILYIPQNYQYDTEQNGHLDLVKAAEEIGIEVQREERNINATAQTSIGTI 67
OY 171 QTAIGLTERGIVLSAPQIDKLLQTKAGQALGSASIVQNNKAKTVLSGIQSLGSVLA 230
Db 68 QTAIGLTERGIVLSAPQIDKLLQTKAGQALGSASIVQNNKAKTVLSGIQSLGSVLA 127
OY 231 GMDLDEALQNNNOHALKAGLELTNSLIENIANSVKTLDFEGRQISQSGKLNKIGLG 290
Db 128 GMDLDEALQNNNOHALKAGLELTNSLIENIANSVKTLDFEGRQISQSGKLNKIGLG 187
OY 291 TLGDKLNKINGIDKAGLGIDVYISGLSGATAALVLDKRNASTAKVAGFELANQVGN 350
Db 188 TLGDKLNKINGIDKAGLGIDVYISGLSGATAALVLDKRNASTAKVAGFELANQVGN 247
OY 351 TKAVSSYLLAQRVAAGLSTGPAALIASTVSLAISPLAFAGIADKFNHAKSLESYAERF 410
Db 248 TRAVSSYLLAQRVAAGLSTGPAALIASTVSLAISPLAFAGIADKFNHAKSLESYAERF 307
OY 411 KRLGVDGNLLAEYRGSTIDASTATATATATAIAGVSAASA-----
Db 308 KRLGVDGNLLAEYRGSTIDASTATATATATAIAGVSAASAASVIAPIALLVSGIT 367
OY 455 -----
Db 368 GVIITLLQSKQAMFEHVANKIHKKIVEMKNNHGNKTFENGIDARYLANLQDNMKFLIN 427
OY 455 -----
Db 428 LNKELQAEERVAITQOQMNINIGLAGISRLGEKVLGSKAYDAFEQKHAKADLYOLD 487
OY 455 -----
Db 488 SANGIIDVNSNGAKTQHILFRTPLTPGTEHREVRQYTGKEYITKLNINRVDNWKITDG 547
OY 455 -----
Db 548 AASTFEDLTNVQRIEILDNAGNVTKTETKIIAKLGEGDNNVFGSGTTEIDGEGYD 607
OY 455 -----
Db 608 RVHYSRGNYGALTIDATKETEGSYTVNRFVETGKALHEVSTHTALVGNREKIEYRHS 667
OY 455 -----
Db 668 NNQHHAGYTTKDTLKAVEEIIIGTSHNDIFKSGFNDAFNGDGVDTIDGNDNRLFGGK 727
OY 455 -----
Db 728 GDDIILGNGDDFDIGGKGNLLHGKGGDIFVHRKGGNDIITDSGNDKLSFSDSNLK 787

```

```

RESULT 10
US-07-779-171-14
: Sequence 14, Application US/0779171
: GENERAL INFORMATION:
: APPLICANT: Potter, Andrew A.
: APPLICANT: Redmond, Mark J.
: APPLICANT: Hughes, Huw P. A.
: TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Morrison & Foerster
: STREET: 545 Middlefield Road, Suite 200
: CITY: Menlo Park
: STATE: California
: COUNTRY: USA
: ZIP: 94025
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/779,171
: FILING DATE: 19911016

```



```

|||||
Db 68 QTAIGLTERIVISAPOIDLKQTRAGQALGSAESIYQANAKAKTVLSGISVLA 127
231 GMDLDELQNNNOHALAKAGLELFTNSLIENINSYKTTDEFEQOISQFSKQNIKGLG 290
128 GMDLDELQNNNOHALAKAGLELFTNSLIENINSYKTTDEFEQOISQFSKQNIKGLG 187
291 TLGDKLKNIGGLDKAGIGLDIVISGLSGATPAALVLADKNASTAKKYGAGFELANOVVNT 350
188 TLGDKLKNIGGLDKAGIGLDIVISGLSGATPAALVLADKNASTAKKYGAGFELANOVVNT 247
351 TKAVSSYIIAQRVAAGISSTGPAALIASTVSLAISPLAFAGIADKFNHAKSLESTAEERF 410
248 TKAVSSYIIAQRVAAGISSTGPAALIASTVSLAISPLAFAGIADKFNHAKSLESTAEERF 307
411 KKLGYDGNLLAEYORGTGTIDASVTAINFALAAIAGVSAALAA----- 454
308 KKLGYDGNLLAEYORGTGTIDASVTAINFALAAIAGVSAALAAAGSVIASPALLVSGIT 367
455 ----- 454
368 GVISTILQYSKOAMFEHVANKIHNKIVEMEKNNHKNYFENGYDARYLANIDQNMKFLN 427
455 ----- 454
428 LNKELQERVIAITQOQWNNIGDLAGISRLGKVLGSKAYVDAFEGKHAKDKLVOLD 487
455 ----- 454
488 SANGIIDVNSGAKTOHILFRTPLTPGTEHREYOTGKYEYITKLNINRVDSWKITDG 547
455 ----- 454
548 AASSTFDLTNVQRIEELDMAGNVTKETKLIKLGEGDDNVFVSGTTEIDGGRYD 607
455 ----- 454
608 RVHYSKNGALTIDATKETEQGSYTVNREVTGKALHEVTSHTALVGNREKIEYRHS 667
455 ----- 454
668 NNQHAQYITKTLKAVEELITGSHNDFKGSFNDAPNGDGVDTIDGNDGDLFGGK 727
455 ----- 454
728 GDDILDGNGDDFLDGKGNDDLHGKGGDIFVHRKGGNDITDSDGNDKLSFSDSNLK 787
455 ----- 454
455 DLFEKYNHMLVITNSKEKVTIONMFREADPAKEVNTKATKDEKIEITIGONGERITS 514
788 DLFEKYNHMLVITNSKEKVTIONMFREADPAKEVNTKATKDEKIEITIGONGERITS 847
515 KOYDDILAKNGKITODELSKYVDNVELLHKSNNVTNSLKLISVSAPFTSSNDSRVLV 574
848 KOYDDILAKNGKITODELSKYVDNVELLHKSNNVTNSLKLISVSAPFTSSNDSRVLV 907
QY 575 APTSMIDQSLSSLOFARGS 593
Db 908 APTSMIDQSLSSLOFARGS 926

```

RESULT 14
US-08-455-970-10
Sequence 10, Application US/08455970

GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.

APPLICANT: REDMOND, MARK J.

APPLICANT: HUGHES, HOW P.A.

TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING LEUKOTOXIN

TITLE OF INVENTION: CHIMERAS

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: REED & ROBINS
STREET: 285 HAMILTON AVENUE, SUITE 200

```

? CITY: PALO ALTO
? STATE: CALIFORNIA
? COUNTRY: UNITED STATES OF AMERICA
? ZIP: 94301
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/455,970
? FILING DATE: 31-MAY-1995
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/960,932
? FILING DATE: 14-OCT-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: ROBINS, ROBERTA L.
? REGISTRATION NUMBER: 33,208
? REFERENCE/DOCKET NUMBER: 9001-0016.10
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 327-3400
? TELEFAX: (415) 327-3231
? INFORMATION FOR SEQ. ID NO: 10:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 943 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-455-970-10

```

Query Match 59.2%; Score 2126; DB 8; Length 943;

Best Local Similarity 52.6%; Pred. No. 1.3e-162;

Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

```

QY 111 SFRKTKAKKIIIVYIPONYQYDTQNGSLDVLKRAEBLGTEYOREERNNTATQTSIGTI 170
Db 8 SFRKTKAKKIIIVYIPONYQYDTQNGSLDVLKRAEBLGTEYOREERNNTATQTSIGTI 67
QY 171 QTAIGLTERIVISAPOIDLKQTRAGQALGSAESIYQANAKAKTVLSGISVLA 230
Db 68 QTAIGLTERIVISAPOIDLKQTRAGQALGSAESIYQANAKAKTVLSGISVLA 127
QY 231 GMDLDELQNNNOHALAKAGLELFTNSLIENINSYKTTDEFEQOISQFSKQNIKGLG 290
Db 128 GMDLDELQNNNOHALAKAGLELFTNSLIENINSYKTTDEFEQOISQFSKQNIKGLG 187
QY 291 TLGDKLKNIGGLDKAGIGLDIVISGLSGATPAALVLADKNASTAKKYGAGFELANOVVNT 350
Db 188 TLGDKLKNIGGLDKAGIGLDIVISGLSGATPAALVLADKNASTAKKYGAGFELANOVVNT 247
QY 351 TKAVSSYIIAQRVAAGISSTGPAALIASTVSLAISPLAFAGIADKFNHAKSLESTAEERF 410
Db 248 TKAVSSYIIAQRVAAGISSTGPAALIASTVSLAISPLAFAGIADKFNHAKSLESTAEERF 307
QY 411 KKLGYDGNLLAEYORGTGTIDASVTAINFALAAIAGVSAALAA----- 454
Db 308 KKLGYDGNLLAEYORGTGTIDASVTAINFALAAIAGVSAALAAAGSVIASPALLVSGIT 367
QY 455 ----- 454
368 GVISTILQYSKOAMFEHVANKIHNKIVEMEKNNHKNYFENGYDARYLANIDQNMKFLN 427
QY 455 ----- 454
428 LNKELQERVIAITQOQWNNIGDLAGISRLGKVLGSKAYVDAFEGKHAKDKLVOLD 487
QY 455 ----- 454
488 SANGIIDVNSGAKTOHILFRTPLTPGTEHREYOTGKYEYITKLNINRVDSWKITDG 547
QY 455 ----- 454

```



```

Db 548 AASTFDLNNVQIGIELDNAGNVTKTKETKIIAKLGGDDNVFVSGTTEIDGEGYD 607
QY 455 ----- 454
Db 608 RVHYSRGNYGALLTDATKETEGSYTVNRFEVETGALHEVSTHVALGNREKELEYRHS 667
QY 455 ----- 454
Db 668 NNQHHAGYTKTTLAVEEIIIGTSHNDIFKSGKFNDAFNGGCVDTIDGNDGDLFGSK 727
QY 455 ----- 454
Db 728 GDDIILGNGDDFIDGKGNDLLHGKGDDIFVHRKGGDNDIITDSGNDKLSFSDSNLK 787
QY 455 ----- 454
Db 455 DLFEEKVHNLYITNSKKEKVTIOWMFRADPAKEVPYKATKDEKIEIITGONERITS 514
Db 788 DLFEEKVHNLYITNSKKEKVTIOWMFRADPAKEVPYKATKDEKIEIITGONERITS 847
QY 515 KOVDLLIAGNGKITODELSKYVDVYELLKHSKNYNSLDKLSSVSFTSSNDSRNLY 574
Db 848 KOVDLLIAGNGKITODELSKYVDVYELLKHSKNYNSLDKLSSVSFTSSNDSRNLY 907
QY 575 APTSMLOSLSLOFARGS 593
Db 908 APTSMLOSLSLOFARGS 926

```

RESULT 15

US-07-779-171-8

Sequence 8, Application US/07779171

GENERAL INFORMATION:

```

APPLICANT: Potter, Andrew A.
APPLICANT: Redmond, Mark J.
APPLICANT: Hughes, Huw P.A.
TITLE OF INVENTION: ENHANCED IMMUNOGENICITY USING
TITLE OF INVENTION: PASTURELLA HEAMOLYTICA LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Morrison & Foerster
STREET: 545 Middlefield Road, Suite 200
CITY: Menlo Park
STATE: California
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07779,171
FILING DATE: 19911016
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Robins, Roberta L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 29310-2001600
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-327-7250
TELEFAX: 415-327-2951
TELEX: 706141
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 951 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-779-171-8

```

Query Match 59.2%; Score 2126; DB 3; Length 951;
 Best Local Similarity 52.6%; Pred. No. 1,3e-162;

```

Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;
QY 111 SFPKTKAKKIIILYIPQNYQYDTEQNGLODLYKAAEELGIEVQREERNRIATAQTSIGTI 170
Db 8 SFPKTKAKKIIILYIPQNYQYDTEQNGLODLYKAAEELGIEVQREERNRIATAQTSIGTI 67
QY 171 QTAIGLTERGIYLSAPQIDKLLQKTRKAGALGASASIYONAKKATVYSGISIGSYLA 230
Db 68 QTAIGLTERGIYLSAPQIDKLLQKTRKAGALGASASIYONAKKATVYSGISIGSYLA 127
QY 231 GMDLDEALONNSNOHALAKAGLELTNSLENIANSVKTLDEFEQIISQFSKLONIKIG 290
Db 128 GMDLDEALONNSNOHALAKAGLELTNSLENIANSVKTLDEFEQIISQFSKLONIKIG 187
QY 291 TLGDKLKNIGGLDKAGLGIDVYISGLSGATTAALVADKNASTAKKAVGAFELANOVGNI 350
Db 188 TLGDKLKNIGGLDKAGLGIDVYISGLSGATTAALVADKNASTAKKAVGAFELANOVGNI 247
QY 351 TKAVSSYIIAORVAAGLSSTGPVALLIATSVSLATISPLAFAGIADKFNHAKSLESYAEER 410
Db 248 TKAVSSYIIAORVAAGLSSTGPVALLIATSVSLATISPLAFAGIADKFNHAKSLESYAEER 307
QY 411 KRLGYDGNLAEYQRTGTTIDASVTAINTALAAIAGVSAANA----- 454
Db 308 KRLGYDGNLAEYQRTGTTIDASVTAINTALAAIAGVSAANAAGSVIATALLVSGIT 367
QY 455 ----- 454
Db 368 GVISTIIQYSKQAMFEHVANKIHNKIVWEKNHNGKNYFENGYDARYLANLODNKKFLIN 427
QY 455 ----- 454
Db 428 LNKELQAEKVIAITQOQNDNINIGDLAGISRLGEKVLGSKAVYDAFEENKHIKAKLVOLD 487
QY 455 ----- 454
Db 488 SANGIIVDSNGSKACTQHILFRTPLLTPTGTEHREVRQYTGKYEYITKLNINRYSKITTG 547
QY 455 ----- 454
Db 548 AASTFDLNNVQIGIELDNAGNVTKTKETKIIAKLGGDDNVFVSGTTEIDGEGYD 607
QY 455 ----- 454
Db 608 RVHYSRGNYGALLTDATKETEGSYTVNRFEVETGALHEVSTHVALGNREKELEYRHS 667
QY 455 ----- 454
Db 668 NNQHHAGYTKTTLAVEEIIIGTSHNDIFKSGKFNDAFNGGCVDTIDGNDGDLFGSK 727
QY 455 ----- 454
Db 728 GDDIILGNGDDFIDGKGNDLLHGKGDDIFVHRKGGDNDIITDSGNDKLSFSDSNLK 787
QY 455 DLFEEKVHNLYITNSKKEKVTIOWMFRADPAKEVPYKATKDEKIEIITGONERITS 514
Db 788 DLFEEKVHNLYITNSKKEKVTIOWMFRADPAKEVPYKATKDEKIEIITGONERITS 847
QY 515 KOVDLLIAGNGKITODELSKYVDVYELLKHSKNYNSLDKLSSVSFTSSNDSRNLY 574
Db 848 KOVDLLIAGNGKITODELSKYVDVYELLKHSKNYNSLDKLSSVSFTSSNDSRNLY 907
QY 575 APTSMLOSLSLOFARGS 593
Db 908 APTSMLOSLSLOFARGS 926

```

Search completed: March 2, 2001, 10:59:14
 Job time: 384 sec

THIS PAGE BLANK (USPTO)

04 of: US-09-306-689-12 to: A_Geneseq_36.* out_format : pfs
 Date: Mar 2, 2001 10:31 AM

About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODE=framed+npz.model -DEV=xip
 -O=cgcn1_1/USPRO.spool/US09306689/runat_02032001_102824_9522/app-query.fasta.1.2389
 -DB=A_Geneseq_36 -QFMT=fastan -SUFFIX=rag -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blonsum62
 -TRANS=human4.csi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM-ext -MINLEN=0 -MAXLEN=2000000000
 -USPR=US09306689_@CGCN1_1_187 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPPX
 -WAIT -THREADS=1

Search information block:

Query: US-09-306-689-12
 Query Length: 2088
 Database: A_Geneseq_36.*
 Database sequences: 268485
 Database length: 34193795
 Search time (sec): 109.740000

score_list:

Sequence	Strid	Orig	ZScore	EScore	len	Documentation	8.7e-27	758	1 En
/SIDS1/gcgdata/geneseq/geneseq/AI1996.DAT:R1187 + 380.50					576.12	5.7e-25	323	1	PI

seq_name: /SIDS1/gcgdata/geneseq/geneseq/AI1996.DAT:R86998

seq_documentation block:

ID R86998 standard; Protein: 758 AA.

AC R86998;

DT 04-JUL-1996 (first entry)

DE Enterohaemorrhagic E.coli hlyA gene product.

KW Enterohaemorrhagic Escherichia coli; virulent; EHEC; O157:H7 serotype;

KM detection; probe; primer; hlyA gene; enterohaemorrhagic colitis;

KX haemolytic uremic syndrome; mesenteric adenitis.

OS Escherichia coli (enterohaemorrhagic).

XX US5475098-A.

XX 12-DEC-1995.

XX 14-JUN-1994; 94US-0258188.

XX 14-JUN-1994; 94US-0258188.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Hall RH, Xu JG;

XX WPI; 1996-048546/05.

XX DR N-PSDB; T08098.

XX Enterohaemorrhagic E. coli (EHEC) nucleic acid sequences - useful
 PT for probe and primer design for sensitive and specific detection of
 PT EHEC
 PS Claim 1; Columns 37-42; 32pp; English.

XX Enterohaemorrhagic E.coli (EHEC) associated with enterohaemorrhagic
 CC colitis, haemolytic uremic syndrome and mesenteric adenitis have
 CC been found to carry a hlyA gene and a hlyB gene, separated by an
 CC intergenic region. The hly genes and the intergenic region are
 CC absent from bacteria not associated with these diseases and so
 CC provide a useful target for detecting EHEC pathogens, esp. O157:H7
 CC serotype E.coli. The present sequence is that of the protein
 CC encoded by the EHEC hlyA gene.

XX Sequence 758 AA;

alignment_scores:

Quality:	403.00	Length:	395
Ratio:	1.807 <td>Gaps:</td> <td>11</td>	Gaps:	11
Percent Similarity:	56.456	Percent Identity:	30.633

alignment block:

US-09-306-689-12 x R86998

Align seg 1/1 to: R86998 from: 1 to: 758

964	GCACCTGTACTTGCAGATAAAATGCTTCACAGCTAAAGAGTGGTGC	1013
1	SerPheIleLeuGlyAsnSerAlaHisThrGlyThrLysAlaAlaI	17
1014	GGGTTTGAATTGGCAACCAAGTTGGTAATATTACCAAGCCGTTT	1063
17	aclyIleGluLeuThrThrGlnValLeuGlyAsnValGlyLysAlaValS	34
1064	CTTCTTACATTTTATAGCCCAAGTGTGCAGAGCTTATCTCACTGGG	1113
34	etGlnIlyrIleLeuAlaGlnArgMetAlaGlnGlyLeuSerThrAla	50
1114	CCTGTTGGCTCTTAAATGCTTCTACTGTTCTGTCGATTAAGCCATT	1163
51	AlaSerAlaGlyLeuIleThrSerAlaValMetLeuAlaIleSerPro	67
1164	AGCATTTGCCGGTATTCGCGATTAATTTAATCATGCAAAAGTTAGAGA	1213
67	uSerPheLeuAlaIleAlaAspLysPheGluArgAlaLysGlnLeuG	84
1214	GTTATGCCGAGCGCTTAAATAATAGCTATGACGCAATATATTATTA	1263
84	etTySerGlnIlyrPheLysLysLeuAsnTyGlnLysAlaLeuLeu	100
1264	GCAGATATATCAGCGGAGGAGGAGCTATGATGATGATGATGATGAT	1313
101	AlaAlaPheHisLysGlnThrGlyAlaIleAspAlaAlaLeuThrTrl	117
1314	TAAATACGATTTGCCGCTATTTGCTGGTGCTGTCTGCTGTCGACGC	1362
117	eAsnThrValLeuSerSerValSerAlaGlyValSerAlaAlaSerSera	134
1362	1362
134	leSerLeuIleGlyAlaProIleSerMetLeuValSerAlaLeuThrGly	150
1363GATTTAACATTTGAAAGT	1382
151	ThrIleSerGlyIleLeuGlnIlyrAlaSerLysGlnAlaMetPheGlnHis	167
1383	TAAACATAATCTTGTATCAGCAATAGCAAAAGAGAAAGTGCACCTTC	1432
167	lAlaGlnLysPheAlaAlaArgIleAsnGlnIlyrGlnLysGlnHisGlyL	184
1433	AAACGTGTTCCGAGAG	1449
184	YasnTyrPheGlnLysGlnIlyrAspAlaArgHisAlaAlaPheLeuGln	200
1450GCTGATTTTGTCTAAAGAGTCCTAATTATTA	1481

1663AGTAATGCTATCTGAAGGCAATTACGCG	16959
1119	LYALAAlaValGlyAlaAlaAlaGlyAlaGlyIleHisGlyTyrThrSer	135
1696	TCTAATGATTCGAGAAAT.....GTATT	1718
136	AsmSerAlaGlyAlaThrLeuGlyAlaGlyLeuAlaAlaGlyLeuValG	152
1719	AGTGGCTCCAACTTCATATCTTGGATCA.....AGTTATCTT	17566
152	YMetAlaAlaAspAlaMetValGluAspValAsnTyrThrMetIleThrA	169
1757	CTCTCAATTTGCT.....	17707
169	SPVALGlnIleAlaGluAlaGlyThrIlySAlaThrValThrThAspAsnVal	185
1771AGGGATCTCAGCA	1784
186	AlaAlaLeuAlaGlnIlyTyrSerGlyAlaIlySileGlnThrSerGlnI	202
1785	TTGGAGACTAGGCGCTCGCGGCGGCAAGCGGTCTCAAGATTGGAGTACG	1834
202	STRPSTYRGLYLeuAlaArgProGly.....GlnHisTyrSerTyrG	216
1835	GCCTGGCTCGCGGAGCTCTACCAACGATTGGAGCTACGCGCTCGCGCT	1884
216	LYLeuAlaArgProGly.....GlnHisTyrSerTyrGlyLeuAlaArg	229
1885	GGCAGCGGATGCCAAGATTGGAGCTACGCGCTCGCGGATGATCTCA	1934
230	Gly.....GlnHisTyrSerTyrGlyLeuAlaArgProGly.....G	241
1935	GCATTGGAGCTACGCGCTCGCGGCGGCAAGCGGTCTCAAGATTGGAGCT	1984
241	HNHSTRPSTYRGLYLeuAlaArgProGly.....GlnHisTyrSerT	255
1985	ACGGCTCGCTCGGAGTGGAGCTTACCCAGATTTGGAGCTACGCGCTCGC	2034
255	YRGLYLeuAlaArgProGly.....GlnHisTyrSerTyrGlyLeuAla	268
2035	CGTGAGCAGCGGATGCCAAGATTGGAGCTACGCGGCTCGCGCGGT	2079
269	PROGly.....GlnHisTyrSerTyrGlyLeuAlaArgProGly	280

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:54:37 ; Search time 47.48 Seconds
(without alignments)
262.851 Million cell updates/sec

Title: US-09-306-689-13

Perfect score: 3591

Sequence: 1 MATVIDRSGQMSYGLRPGSG.....GLRPGSGSDMSYGLRPGSG 695

Scoring table:

BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents AA: *
2: /cgn2_6/ptodata/2/1aa/5A.COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6.COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/PCYTUS.COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/Backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	3591	100.0	699	2	US-08-694-865-16 Sequence 16, Appl
2	3591	100.0	699	3	US-09-124-491-16 Sequence 16, Appl
3	2642.5	73.6	544	1	US-08-387-156-10 Sequence 10, Appl
4	2642.5	73.6	544	2	US-08-694-865-10 Sequence 10, Appl
5	2642.5	73.6	544	2	US-08-878-748-10 Sequence 10, Appl
6	2642.5	73.6	544	3	US-09-124-491-10 Sequence 10, Appl
7	2426	67.6	977	2	US-08-387-156-8 Sequence 8, Appl
8	2426	67.6	977	2	US-08-694-865-8 Sequence 8, Appl
9	2426	67.6	977	2	US-08-878-748-8 Sequence 8, Appl
10	2426	67.6	977	3	US-09-124-491-8 Sequence 8, Appl
11	2189	61.0	936	1	US-08-455-970A-12 Sequence 12, Appl
12	2131	59.3	1069	1	US-07-777-715-9 Sequence 9, Appl
13	2131	59.3	1069	1	US-08-170-126-4 Sequence 4, Appl
14	2131	59.3	1069	3	US-08-954-418-4 Sequence 4, Appl
15	2126	59.2	926	1	US-07-908-253-2 Sequence 2, Appl
16	2126	59.2	926	1	US-08-455-970A-2 Sequence 2, Appl
17	2126	59.2	926	1	US-08-387-156-6 Sequence 6, Appl
18	2126	59.2	926	2	US-08-694-865-6 Sequence 6, Appl
19	2126	59.2	926	2	US-08-878-748-6 Sequence 6, Appl
20	2126	59.2	926	2	US-08-535-837-2 Sequence 2, Appl
21	2126	59.2	926	3	US-09-124-491-6 Sequence 6, Appl
22	2126	59.2	926	5	5476657-3 Patent No. 5476657
23	2126	59.2	943	1	US-08-455-970A-10 Sequence 10, Appl
24	2126	59.2	951	1	US-08-455-970A-14 Sequence 14, Appl
25	2117.5	59.0	1098	1	US-07-777-715-7 Sequence 7, Appl
26	2117.5	59.0	1098	1	US-08-170-126-2 Sequence 2, Appl
27	2117.5	59.0	1098	3	US-08-954-418-2 Sequence 2, Appl
28	2116	58.9	924	3	US-08-619-812-8 Sequence 8, Appl

29	1650.5	46.0	1403	1	US-07-908-253-3 Sequence 3, Appl
30	1650.5	46.0	1403	2	US-08-694-865-17 Sequence 17, Appl
31	1650.5	46.0	1403	2	US-08-535-837-3 Sequence 3, Appl
32	1650.5	46.0	1403	3	US-09-124-491-17 Sequence 17, Appl
33	1644.5	45.8	1334	5	5476657-1 Patent No. 5476657
34	1524	42.4	934	1	US-08-215-805A-80 Sequence 80, Appl
35	1264.5	35.2	956	3	US-08-772-270A-8 Sequence 8, Appl
36	938.5	26.1	1049	3	US-08-772-270A-11 Sequence 11, Appl
37	938.5	26.1	1244	4	PCT-US93-10500-2 Sequence 2, Appl
38	848.5	23.6	1022	3	US-08-772-270A-2 Sequence 2, Appl
39	403	11.2	758	1	US-08-258-188-2 Sequence 2, Appl
40	403	11.2	758	1	US-08-526-813-2 Sequence 2, Appl
41	403	11.2	758	4	PCT-US95-08554-2 Sequence 2, Appl
42	368	10.2	84	1	US-07-690-983D-47 Sequence 47, Appl
43	290	8.1	49	1	US-08-387-156-4 Sequence 4, Appl
44	290	8.1	49	2	US-08-694-865-4 Sequence 4, Appl
45	290	8.1	49	2	US-08-878-748-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-694-865-16
Sequence 16, Application US/08694865
Patent No. 5837268
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANN, JOHN G.
TITLE OF INVENTION: GPRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,865
FILING DATE: 09-AUG-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 699 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-865-16

Query Match 100.0%; Score 3591; DB 2; Length 699;
Best Local Similarity 100.0%; Pred. No. 2,9e+263;
Matches 695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATVIDRSGQMSYGLRPGSGSDMSYGLRPGSGSDMSYGLRPGSG 60
DB 1 MATVIDRSGQMSYGLRPGSGSDMSYGLRPGSGSDMSYGLRPGSG 60
QY 61 HMSYGLRPGSGSDMSYGLRPGSGSDMSYGLRPGSGSDMSYGLRPGSGPPTGAKKI 120
|||||

Db 128 GMDLDELALONNSQNALAKAGLELTNSLIENIANSVKTLDEFGEOISQFGSKLQNIKIGL 187
QY 291 TLGDKLKNIGGLDKAGLDIVISGLISGATPAALVLADKNASTAKKYGAGFELANOVVGN 350
Db 188 TLGDKLKNIGGLDKAGLDIVISGLISGATPAALVLADKNASTAKKYGAGFELANOVVGN 247
QY 351 TKAVSSYIIAORVAAGLSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLESYAERF 410
Db 248 TKAVSSYIIAORVAAGLSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLESYAERF 307
QY 411 KRLGIDGNILAEYQGTGTIDASVTAINPALAIAIGVSAANA---DLTFEYKHNLYI 467
Db 308 KRLGIDGNILAEYQGTGTIDASVTAINPALAIAIGVSAANA---DLTFEYKHNLYI 367
QY 468 TNSKKEKVTIOMNFRADPAKEVPNKATKDEKIEEIIIGONGERITISKOVDDLIAGNGK 527
Db 368 TNSKKEKVTIOMNFRADPAKEVPNKATKDEKIEEIIIGONGERITISKOVDDLIAGNGK 427
QY 528 ITODELSKVVDNELKHSKNVTNSLDKLISVSATSSNDSRNVLVAPTSMLDQSLSSL 587
Db 428 ITODELSKVVDNELKHSKNVTNSLDKLISVSATSSNDSRNVLVAPTSMLDQSLSSL 487
QY 588 QFARGSQHSYGLRPGSSQDMSYGLRPGSSQHSYGLRPGSSQDMSYGLRPGSS 644
Db 488 QFARGSQHSYGLRPGSSQDMSYGLRPGSSQHSYGLRPGSSQDMSYGLRPGSS 544

RESULT 5

US-08-878-748-10
; Sequence 10, Application US/08878748
; Patent No. 5969126
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HOW P. A.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,748
; FILING DATE: 19-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/387,156
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016, 21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-08-878-748-10

Query Match 73.6%; Score 2642.5; DB 2; Length 544;
Best Local Similarity 99.4%; Pred. No. 8; 9e-192;
Matches 534; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 111 SPFKTAKKIIITYIPONYQYDFQNGLODLVKAARELGIEVOREERNNTATQTSIGTI 170
Db 8 SPFKTAKKIIITYIPONYQYDFQNGLODLVKAARELGIEVOREERNNTATQTSIGTI 67
QY 171 QTAIGLTERGIYLSAPQIDKLQKTAKGALSAESIYONANKAKTVLSIGIOSISVLA 230
Db 68 QTAIGLTERGIYLSAPQIDKLQKTAKGALSAESIYONANKAKTVLSIGIOSISVLA 127
QY 231 GMDLDELALONNSQNALAKAGLELTNSLIENIANSVKTLDEFGEOISQFGSKLQNIKIGL 290
Db 128 GMDLDELALONNSQNALAKAGLELTNSLIENIANSVKTLDEFGEOISQFGSKLQNIKIGL 187
QY 291 TLGDKLKNIGGLDKAGLDIVISGLISGATPAALVLADKNASTAKKYGAGFELANOVVGN 350
Db 188 TLGDKLKNIGGLDKAGLDIVISGLISGATPAALVLADKNASTAKKYGAGFELANOVVGN 247
QY 351 TKAVSSYIIAORVAAGLSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLESYAERF 410
Db 248 TKAVSSYIIAORVAAGLSTGPVAALIASTVSLAISPLAFAGIADKFNHAKSLESYAERF 307
QY 411 KRLGIDGNILAEYQGTGTIDASVTAINPALAIAIGVSAANA---DLTFEYKHNLYI 467
Db 308 KRLGIDGNILAEYQGTGTIDASVTAINPALAIAIGVSAANA---DLTFEYKHNLYI 367
QY 468 TNSKKEKVTIOMNFRADPAKEVPNKATKDEKIEEIIIGONGERITISKOVDDLIAGNGK 527
Db 368 TNSKKEKVTIOMNFRADPAKEVPNKATKDEKIEEIIIGONGERITISKOVDDLIAGNGK 427
QY 528 ITODELSKVVDNELKHSKNVTNSLDKLISVSATSSNDSRNVLVAPTSMLDQSLSSL 587
Db 428 ITODELSKVVDNELKHSKNVTNSLDKLISVSATSSNDSRNVLVAPTSMLDQSLSSL 487
QY 588 QFARGSQHSYGLRPGSSQDMSYGLRPGSSQHSYGLRPGSSQDMSYGLRPGSS 644
Db 488 QFARGSQHSYGLRPGSSQDMSYGLRPGSSQHSYGLRPGSSQDMSYGLRPGSS 544

RESULT 6

US-09-124-491-10
; Sequence 10, Application US/09124491
; Patent No. 6022960
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,491
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE: 09-AUG-1996

APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491-10

Query Match 73.6%; Score 2642.5; DB 3; Length 544;
Best Local Similarity 99.4%; Pred. No. 8.9e-192;
Matches 534; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 111 SEPKTGAKKIIILYIPONTQYDTEQNGIQLDVLKAAEELGIEVOREERNINATAQTSIGTI 170
DB 8 SEPKTGAKKIIILYIPONTQYDTEQNGIQLDVLKAAEELGIEVOREERNINATAQTSIGTI 67
OY 171 QTAIGLTERGIVLSAPQIDKLQKTKAGQALGSAESIYONANKAKTVLSGISIIGSVLA 230
DB 68 QTAIGLTERGIVLSAPQIDKLQKTKAGQALGSAESIYONANKAKTVLSGISIIGSVLA 127
OY 231 GMDLDEALONNSNOHALKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQNIKIGL 290
DB 128 GMDLDEALONNSNOHALKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQNIKIGL 187
OY 291 TLGDKLKNIGGLDKAGLDIVISGLSGATAALVLADKNASTAKKVGAFELANQVGN 350
DB 188 TLGDKLKNIGGLDKAGLDIVISGLSGATAALVLADKNASTAKKVGAFELANQVGN 247
OY 351 TKAVSSYLLAORVAAAGLSSTGPVALLASTVSLAISPLAFAGIADKFHNAKSLESYAERF 410
DB 248 TKAVSSYLLAORVAAAGLSSTGPVALLASTVSLAISPLAFAGIADKFHNAKSLESYAERF 307
OY 411 KKLGYDGNLLAEYORGTGTIDASTATNTALAAIAGVSAAAA---DLTEPKVKNHNYI 467
DB 308 KKLGYDGNLLAEYORGTGTIDASTATNTALAAIAGVSAAAA---DLTEPKVKNHNYI 367
OY 468 TNSKKEKVTIOWMFREADFAKEVPYKATKDEKIEIIGONGERTTSKQVODLIKGNK 527
DB 368 TNSKKEKVTIOWMFREADFAKEVPYKATKDEKIEIIGONGERTTSKQVODLIKGNK 427
OY 528 ITQDLSLVVNYELKLSKNTNSLDKLISVSFAFTSSNDSRNVLPVPTSLDQSLSSL 587
DB 428 ITQDLSLVVNYELKLSKNTNSLDKLISVSFAFTSSNDSRNVLPVPTSLDQSLSSL 487
OY 588 QFARSSQHSYGLRPGSSQDMSYGLRPGSSQHSYGLRPGSSQDMSYGLRPGSS 644
DB 488 QFARSSQHSYGLRPGSSQDMSYGLRPGSSQHSYGLRPGSSQDMSYGLRPGSS 544

RESULT 7
US-08-387-156-8
Sequence 8, Application US/08387156
Patent No. 5723129
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HUW P. A.

TITLE OF INVENTION: GHRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESSES:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/387,156
FILING DATE: 10-FEB-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 977 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-387-156-8

Query Match 67.6%; Score 2426; DB 1; Length 977;
Best Local Similarity 55.1%; Pred. No. 4.9e-175;
Matches 534; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

OY 111 SEPKTGAKKIIILYIPONTQYDTEQNGIQLDVLKAAEELGIEVOREERNINATAQTSIGTI 170
DB 8 SEPKTGAKKIIILYIPONTQYDTEQNGIQLDVLKAAEELGIEVOREERNINATAQTSIGTI 67
OY 171 QTAIGLTERGIVLSAPQIDKLQKTKAGQALGSAESIYONANKAKTVLSGISIIGSVLA 230
DB 68 QTAIGLTERGIVLSAPQIDKLQKTKAGQALGSAESIYONANKAKTVLSGISIIGSVLA 127
OY 231 GMDLDEALONNSNOHALKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQNIKIGL 290
DB 128 GMDLDEALONNSNOHALKAGLELTNSLIENIANSVKTLDEFGEQISQFGSKLQNIKIGL 187
OY 291 TLGDKLKNIGGLDKAGLDIVISGLSGATAALVLADKNASTAKKVGAFELANQVGN 350
DB 188 TLGDKLKNIGGLDKAGLDIVISGLSGATAALVLADKNASTAKKVGAFELANQVGN 247
OY 351 TKAVSSYLLAORVAAAGLSSTGPVALLASTVSLAISPLAFAGIADKFHNAKSLESYAERF 410
DB 248 TKAVSSYLLAORVAAAGLSSTGPVALLASTVSLAISPLAFAGIADKFHNAKSLESYAERF 307
OY 411 KKLGYDGNLLAEYORGTGTIDASTATNTALAAIAGVSAAAA-----+ 454
DB 308 KKLGYDGNLLAEYORGTGTIDASTATNTALAAIAGVSAAAA-----+ 367
OY 455 ----- 454
DB 368 GVSTIIQYSKQMEHVAHNRKIHNVEMERNNHKNYFENGYDARYLANLDNMKFLAN 427
OY 455 ----- 454

|||||
Db 908 APTSMLDOSLSLQFARSGSHWSYGLRPGSGSDMSYGLRPGSGSHWSYGLRPGSGSD 967
QY 635 WSYGLRPGGS 644
|||||
Db 968 WSYGLRPGGS 977

RESULT 9

US-08-878-748-8
Sequence 8, Application US/08878748
Patent No. 5969126
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HOW P. A.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,748
FILING DATE: 19-JUN-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 977 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-748-8

Query Match 67.6%; Score 2426; DB 2; Length 977;
Best Local Similarity 55.1%; Pred. No. 4.9e-175;
Matches 534; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

QY 111 SFPKTKGAKKILYIPONTQYDTGCGNGIQLDYKAAEEIGIEYQREERNINATAGTSLGTI 170
|||||
Db 8 SFPKTKGAKKILYIPONTQYDTGCGNGIQLDYKAAEEIGIEYQREERNINATAGTSLGTI 67
QY 171 QTAIGLTERGIVLSAPQIDKLQKTRKAGALGASAEIVQNNANKAKTVSGIOSIIGSVILA 230
|||||
Db 68 QTAIGLTERGIVLSAPQIDKLQKTRKAGALGASAEIVQNNANKAKTVSGIOSIIGSVILA 127
QY 231 GMDLDEALQNNNOHALAKAGIELTNSLIENIANSVKTLDGEQGISQFGSKLQNIKIGLG 290
|||||
Db 128 GMDLDEALQNNNOHALAKAGIELTNSLIENIANSVKTLDGEQGISQFGSKLQNIKIGLG 187

QY 291 TLGDKLKNIGGLDKAGLGLDVIISGLLSGATAALVLDKNASTAKKVGAGFELANOVGNI 350
|||||
Db 188 TLGDKLKNIGGLDKAGLGLDVIISGLLSGATAALVLDKNASTAKKVGAGFELANOVGNI 247
QY 351 TKAVSSYILAQRVAGLSTGTPVAALIASVSLAISPLAFAGIADKFNHAKSLESYAERF 410
|||||
Db 248 TKAVSSYILAQRVAGLSTGTPVAALIASVSLAISPLAFAGIADKFNHAKSLESYAERF 307
QY 411 KKLGYDGNLLAEYORGCTGTIDASYTAINTALAAIAGVSAAAA----- 454
|||||
Db 308 KKLGYDGNLLAEYORGCTGTIDASYTAINTALAAIAGVSAAAAASVIAISPILLVSGIT 367
QY 455 ----- 454
Db 368 GVIISTIIQYSQAMFEHVANKINKIVEMKNNHKNYFENGYDARYLANIQQNNKFLIN 427
QY 455 ----- 454
Db 428 LNKELQAEVLAITQQQMDNNIGDLAGISRLGEKYLISKRAYVDAFEBSKHIAKDLVOLD 487
QY 455 ----- 454
Db 488 SANGIIDVNSGKAKTOHILFRTPLTPGTEHREVRGTGYEYITKLNINRVDSWKITDG 547
QY 455 ----- 454
Db 548 AASTFEDLTNNVQRIEILDNAGNVTKREKRIIAKLGEQDNNFVGSQTTEIDGEGYD 607
QY 455 ----- 454
Db 608 RVHYSRGVAGLITDATETEDQSYTVNRPVETGALHEVSTHTALVGNREKIEYRHS 667
QY 455 ----- 454
Db 668 NNOHHAGYTTKDLKAVEIIGTSHNDIFKSGKFNDAPFNGDGYDTIDGNDGNRLFGGK 727
QY 455 ----- 454
Db 728 GDDILIDGNGDDFIDGKGNDLHGCGKDDIFVHRKGDNDIITDSQDNKLSFSDSNLK 787
QY 455 DTFEKKVKNLVTNSKKKQYTIQNMFEADFAKVPYKATKDEKIEIIGONERITS 514
|||||
Db 788 DTFEKKVKNLVTNSKKKQYTIQNMFEADFAKVPYKATKDEKIEIIGONERITS 847
QY 515 KOVDDLAKNGKITQDELISKVVDNYELLKHSKNVTNSLDKLISVSASFSSNDSRNLY 574
|||||
Db 848 KOVDDLAKNGKITQDELISKVVDNYELLKHSKNVTNSLDKLISVSASFSSNDSRNLY 907
QY 575 APTSMLDOSLSLQFARSGSHWSYGLRPGSGSDMSYGLRPGSGSHWSYGLRPGSGSD 634
|||||
Db 908 APTSMLDOSLSLQFARSGSHWSYGLRPGSGSDMSYGLRPGSGSHWSYGLRPGSGSD 967
QY 635 WSYGLRPGGS 644
|||||
Db 968 WSYGLRPGGS 977

RESULT 10
US-09-124-491-8
Sequence 8, Application US/09124491
Patent No. 6022960

GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANN, JOHN G.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA

TELEFAX: (415) 327-3231
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 936 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-455-970A-12

Query Match 61.0%; Score 2189; DB 1; Length 936;
 Best Local Similarity 53.1%; Pred. No. 3,7e-157;
 Matches 493; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

QY 111 SPPKGAKKIILYIPONYOYDEONGLODLYKAEELGIEVQREERNINIAAOTSLGTI 170
 DB 8 SPPKGAKKIILYIPONYOYDEONGLODLYKAEELGIEVQREERNINIAAOTSLGTI 67
 QY 171 QTAIGLTERGIVLSAPQIDKLLQKTKAGALGSASISIVQANKAKTVLSGISIILGSVLA 230
 DB 68 QTAIGLTERGIVLSAPQIDKLLQKTKAGALGSASISIVQANKAKTVLSGISIILGSVLA 127
 QY 231 GMDLDEALQNNNSNOHALAKAGLELTNSLIENIANSVKTLDEFGEOISQFGSKLQNIKIGL 290
 DB 128 GMDLDEALQNNNSNOHALAKAGLELTNSLIENIANSVKTLDEFGEOISQFGSKLQNIKIGL 187
 QY 291 TLGDKLKNIGIGLDKRGGLDLYISGLLSGATATLVADKNASTAKKVGAGFELANOVGNI 350
 DB 188 TLGDKLKNIGIGLDKRGGLDLYISGLLSGATATLVADKNASTAKKVGAGFELANOVGNI 247
 QY 351 TKAVSSYLLAORVAAGLSSTGPVALLIATSVSLAISPLAFAGIADKFHNAKSLESYAERF 410
 DB 248 TKAVSSYLLAORVAAGLSSTGPVALLIATSVSLAISPLAFAGIADKFHNAKSLESYAERF 307
 QY 411 KKLGYDGNLLAEYQRGTTIDASVTATNTALAAIAGVSAASA----- 454
 DB 308 KKLGYDGNLLAEYQRGTTIDASVTATNTALAAIAGVSAASAASYLALVSGIT 367
 QY 455 ----- 454
 DB 368 GVISTIILOYSKOAMFEHVANKIHKKIVEMKNNHKNYFENGYDARYLANQDNKKFLIN 427
 QY 455 ----- 454
 DB 428 LNKELQARVIAITQQQKNDNIGDLAGISRLGEKYLSGKAYVDAREBKHKIKADLVQLD 487
 QY 455 ----- 454
 DB 488 SANGIIDVNSGKAKTQHILFRTPLLPTEHREHREVOTGKEYITKLNINRYDSMKITDG 547
 QY 455 ----- 454
 DB 548 AASSTFDLTNNVQRIIGIELDNAGNVTKREKIIAKLCEGDNDVEVSGTTEIDGEGYD 607
 QY 455 ----- 454
 DB 608 RVHYSRGNYGALTIDATKETEEOGSYTVNRPEVETGKALHEVYSTHTALVGNREKIEYRHS 667
 QY 455 ----- 454
 DB 668 NNQHHAGYYTKDTLKAVEIIGTSHNDIFKSGKFNDAFNGDGVDTIDGNDGNDRLFSGK 727
 QY 455 ----- 454
 DB 728 GDDIIDGNGDDEFIDGKGNDLHGKGGDIFVHRKGGGNDIITDSNDKLSFSDSNLK 787
 QY 455 ----- 454
 DB 788 DTFEKKVHNLYITNSKREKVTIQMFPREADPAKEVPYKATKDKIEIIGONCERITTS 514
 QY 455 ----- 454
 DB 848 KOVDLLIAKGNCKITQDELISKVVDNYELLKHSKNVTNSLKLISVSASFSSNDSRNLYV 574
 DB 848 KOVDLLIAKGNCKITQDELISKVVDNYELLKHSKNVTNSLKLISVSASFSSNDSRNLYV 907

QY 575 APTSMUDOSLSSLOFARGSOHWSYGLRPG 603
 DB 908 APTSMUDOSLSSLOFARGSOHWSYGLRPG 936

RESULT 12 US-07-777-715-9

Sequence 9, Application US/07777715

Patent No. 5273889

GENERAL INFORMATION:

APPLICANT: Potter, Andrew

APPLICANT: Campos, Manuel

APPLICANT: Hughes, Huw P. A.

TITLE OF INVENTION: CYTOKINE-LEUKOTOXIN GENE FUSIONS AND

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Morrison & Foerster

STREET: 545 Middlefield Road, Suite 200

CITY: Menlo Park

STATE: California

COUNTRY: USA

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/777,715

FILING DATE: 19911016

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Robins, Roberta L.

REGISTRATION NUMBER: 33,208

REFERENCE/DOCKET NUMBER: 29310-2001320

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-327-7250

TELEFAX: 415-327-2951

TELEX: 706141

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 1069 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-777-715-9

Query Match 59.3%; Score 2131; DB 1; Length 1069;
 Best Local Similarity 52.6%; Pred. No. 1.1e-152;
 Matches 484; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

QY 111 SPPKGAKKIILYIPONYOYDEONGLODLYKAEELGIEVQREERNINIAAOTSLGTI 170
 DB 8 SPPKGAKKIILYIPONYOYDEONGLODLYKAEELGIEVQREERNINIAAOTSLGTI 67
 QY 171 QTAIGLTERGIVLSAPQIDKLLQKTKAGALGSASISIVQANKAKTVLSGISIILGSVLA 230
 DB 68 QTAIGLTERGIVLSAPQIDKLLQKTKAGALGSASISIVQANKAKTVLSGISIILGSVLA 127
 QY 231 GMDLDEALQNNNSNOHALAKAGLELTNSLIENIANSVKTLDEFGEOISQFGSKLQNIKIGL 290
 DB 128 GMDLDEALQNNNSNOHALAKAGLELTNSLIENIANSVKTLDEFGEOISQFGSKLQNIKIGL 187
 QY 291 TLGDKLKNIGIGLDKRGGLDLYISGLLSGATATLVADKNASTAKKVGAGFELANOVGNI 350
 DB 188 TLGDKLKNIGIGLDKRGGLDLYISGLLSGATATLVADKNASTAKKVGAGFELANOVGNI 247
 QY 351 TKAVSSYLLAORVAAGLSSTGPVALLIATSVSLAISPLAFAGIADKFHNAKSLESYAERF 410
 DB 248 TKAVSSYLLAORVAAGLSSTGPVALLIATSVSLAISPLAFAGIADKFHNAKSLESYAERF 307

```

1 0Y      411 KKGYDGNLLAEYORGCTIDASVTALNTALAIAGVSAAAA-----454
2 0Y      308 KKLGYDGNLLAEYORGCTIDASVTALNTALAIAGVSAAGSVIASPALLVSGIT 367
3 0Y      455 -----454
4 Db      368 GVISITLIQSKQAMPEEHANKIHKKIYEWEMKNNHKNVFPENGYDARYLANLQDNMKFLN 427
5 0Y      455 -----454
6 Db      428 LNKLEAERVIAITQOQWDMNINIGDLAGISRLGEKVLGKAYVDAFEEGKHIAKAKIYOLD 487
7 0Y      455 -----454
8 Db      488 SANGIIDVSNSSGKAKTOHILFRTPLLPCTEHERERVOTGKEIYITTKINRVDWSKITDG 547
9 0Y      455 -----454
10 Db      548 AASSTFDLTNNVVRIGIELDNAGNVTKTKETKILAKLGEDDNVFVSGTTEIDGGEYD 607
11 0Y      455 -----454
12 Db      608 RVHSRGNYGALLTDATKETEGSYVNRFEYTGKALHEVYSTHTALVGNREEKIEYRAS 667
13 0Y      455 -----454
14 Db      668 NNQHHAGIYTTDLKAVEIIGTSHNDIFKSGKFNDAFNGSGDVDTIDGNDGDRLEFGK 727
15 0Y      455 -----454
16 Db      728 GDDILDGNGDDFDIDGKGNDLLHGKGDDIFVHRKGDNDIITDSDGNKLSFSDSNLK 787
17 0Y      455 DLTFEKVXHNLYITNSKKKEKTIONWPREADFAKEVNYKATKKEKIEEIIIGNGEYITS 514
18 Db      788 DLTFEKVXHNLYITNSKKKEKVTIONWPREADFAKEVNYKATKKEKIEEIIIGNGEYITS 847
19 0Y      515 KOVDLILKNGSKTODLSKYVNVYELLKSKVNTNSLDKLISSVAFSSNDSNRVLY 574
20 Db      848 KOVDLILKNGSKTODLSKYVNVYELLKSKVNTNSLDKLISSVAFSSNDSNRVLY 907
21 0Y      575 APTSMLDQSLSSLOFARGSQ 594
22 Db      908 APTSMLDQSLSSLOFARGSQ 927
23
24 RESULT 13
25 US-08-170-126-4
26 : Sequence 4, Application US/08170126
27 : Patent No. 5594107
28 :
29 : GENERAL INFORMATION:
30 :
31 : APPLICANT: POTTER, ANDREW
32 : APPLICANT: CAMPOS, MANUEL
33 : APPLICANT: HUGHES, HOW P. A.
34 : TITLE OF INVENTION: CYTOKINE-CYTOTOXIN GENE FUSIONS AND USES
35 : TITLE OF INVENTION: THEREOF
36 : NUMBER OF SEQUENCES: 6
37 : CORRESPONDENCE ADDRESS:
38 : ADDRESSEE: REED & ROBINS
39 : STREET: 635 BRYANT STREET
40 : CITY: PALO ALTO
41 : STATE: CALIFORNIA
42 : COUNTRY: UNITED STATES OF AMERICA
43 : ZIP: 94301
44 :
45 : COMPUTER READABLE FORM:
46 : MEDIUM TYPE: Floppy disk
47 : COMPUTER: IBM PC compatible
48 : OPERATING SYSTEM: PC-DOS/MS-DOS
49 : SOFTWARE: PatentIn Release #1.0, Version #1.25
50 : CURRENT APPLICATION DATA:
51 : APPLICATION NUMBER: US/08/170,126
52 : FILING DATE: 20-DEC-1993
53 : CLASSIFICATION: 530

```

```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER:   US 07/777,715
3      PRIOR APPLICATION DATA:
4      APPLICATION NUMBER:   US 07/571,301
5      ATTORNEY/AGENT INFORMATION:
6      NAME:  ROBINS, ROBERTA L.
7      REGISTRATION NUMBER:  33,208
8      REFERENCE/DOCKET NUMBER:  9000-0013.21
9      TELECOMMUNICATION INFORMATION:
10     TELEPHONE:  (415) 617-8999
11     TELEFAX:  (415) 347-3231
12     INFORMATION FOR SEQ ID NO:  4:
13     SEQUENCE CHARACTERISTICS:
14     LENGTH: 1069 amino acids
15     TYPE: amino acid
16     TOPOLOGY: linear
17     MOLECULE TYPE: protein
18     US-08-170-126-4

```

Query Match	59.3%;	Score 2131;	DB 1;	length 1069;
Best Local Similarity	52.6%;	Pred. No. 1.1e-152;		
Matches 484;	Conservative	0;	Mismatches 0;	Indels 436; Gaps 1.

OY	111	SEPKGAKKILLYTPONQYOTBOGNGLQDLYKAAEELGTEYVOEKEENNATATQTSIGTI	170
Db	8	SFPKPGAKKILLYTPONQYOTBOGNGLQDLYKAAEELGTEYVOEKEENNATATQTSIGTI	67
OY	171	QTAIGLTERGIVLSAPQIDKLLQKTPKAGQAGSAAESIYVQANAKKATYLSIGTISLSVLA	230
Db	68	QTAIGLTERGIVLSAPQIDKLLQKTPKAGQAGSAAESIYVQANAKKATYLSIGTISLSVLA	127
OY	231	GMDLDEALONNSNOHALAKAGLELTNSLIENIANSVKTLDFEGEIQISQFSKLONTKGLG	230
Db	128	GMDLDEALONNSNOHALAKAGLELTNSLIENIANSVKTLDFEGEIQISQFSKLONTKGLG	187
OY	231	TLGDKLNIIGLIDRAGGLDVIISGLLSGATPAIYLAKNNASTAKKVAGFELNNOVYNI	350
Db	188	TLGDKLNIIGLIDRAGGLDVIISGLLSGATPAIYLAKNNASTAKKVAGFELNNOVYNI	247
OY	351	TKAVSSYTLAORVAAAGLSSTGPVALLIASYSLAISPLAFAGIDRKNHAKSLESYAERE	410
Db	248	TKAVSSYTLAORVAAAGLSSTGPVALLIASYSLAISPLAFAGIDRKNHAKSLESYAERE	307
OY	411	KKLGYDGNLLAEYORGTGTIDASVTAINPALAAIAGVSAAAA	454
Db	308	KKLGYDGNLLAEYORGTGTIDASVTAINPALAAIAGVSAAAA	367
OY	455	-----	454
Db	368	GVTSTIIQYSKOAMFEHVANKIHNKIVEMEKNNHKNYPENGYDARYLANLQDNMKFLN	427
OY	455	-----	454
Db	428	LNKELQARVIAITQOQMDNNIGSLAGISRLGEVLSGKAYVDAFBEGBKHKKADKLQYLD	487
OY	455	-----	454
Db	488	SANGIIDVNSGKAKQHLFPRPLTPGTREHREYQTKGYEYITTKLINRVDSMKITTDG	547
OY	455	-----	454
Db	548	AASSTFDLTNNVORIGIELDAGNAVTKTKETKIIAKLGEDDNVFVSGTTEIDGGEYD	607
OY	455	-----	454
Db	608	RVHSRGNVYKALLTDATKETEGSTYVNRVEYETKALHEVTSHTTALVGNREEKIEYRHS	667
OY	455	-----	454
Db	668	NNQHHAGYYTTDLKAVBEIIGTSHNDIFKSGKRNDAFNGGDGVDTIDGNDGNDRLFGGK	727
OY	455	-----	454


```

Db 728 GDDILDGNGDDFDLGGKNDLHGKGGDIFVHRKGGNDIITDSDGNDKLSFSDSNLK 787
QY 455 DLTEFKVHNLYVITNSKKEKVTIQMWFREADFAKEVPYKATKDEKIEIIGONGERITS 514
Db 788 DLTEFKVHNLYVITNSKKEKVTIQMWFREADFAKEVPYKATKDEKIEIIGONGERITS 847
QY 515 KOVDDLAKGNGKITODELSKVVDNYELKHSKNVTNSLDKLSSVSFAFTSSNDSRNLY 574
Db 848 KOVDDLAKGNGKITODELSKVVDNYELKHSKNVTNSLDKLSSVSFAFTSSNDSRNLY 907
QY 575 APTSMLDQSLSSLOFARGSQ 594
Db 908 APTSMLDQSLSSLOFARGSQ 927

RESULT 14
US-08-954-418-4
; Sequence 4, Application US/08954418
; Patent No. 6096320
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW
; APPLICANT: CAMPOS, MANUEL
; APPLICANT: HUGHES, HOW P. A.
; TITLE OF INVENTION: CYTOKINE-CYTOFOXIN GENE FUSIONS AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,418
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170,126
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/571,301
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9000-0013.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1069 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-954-418-4

Query Match 59.38; Score 2131; DB 3; Length 1069;
Best Local Similarity 52.68; Pred. No. 1,1e-152;
Matches 484; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

QY 111 SEPTGAKKILYIPONQYDTBOGNGLODLYVKAEEIGIEVOREERNINATAQTSIGTI 170
Db 8 SEPTGAKKILYIPONQYDTBOGNGLODLYVKAEEIGIEVOREERNINATAQTSIGTI 67
QY 171 QTAIGLTERGIVLSAPQIDKLQKTKAGQALSAESIYQNNANKAKTVLSIGISLVIA 230

```

```

Db 68 QTAIGLTERGIVLSAPQIDKLQKTKAGQALSAESIYQNNANKAKTVLSIGISLVIA 127
QY 231 GMDLDEALQNNNSQHALAKAGLELTNSLIENIANSVKTLDEFGQDISPGSKLONIKGLG 290
Db 128 GMDLDEALQNNNSQHALAKAGLELTNSLIENIANSVKTLDEFGQDISPGSKLONIKGLG 187
QY 291 TLGDKLNKINGLDRAGLGLDVIYSGLSGATPAALVADNASTAKKVGAGFELANQVGN 350
Db 188 TLGDKLNKINGLDRAGLGLDVIYSGLSGATPAALVADNASTAKKVGAGFELANQVGN 247
QY 351 TKAVSSYTLAQRVAAGLSSTGPVAAALASTVSLASPLAFAGIADKFNHAKSLESYARE 410
Db 248 TKAVSSYTLAQRVAAGLSSTGPVAAALASTVSLASPLAFAGIADKFNHAKSLESYARE 307
QY 411 KKLGYDGNLLAEYORGTTIDASTYTAINTALAIAGVSAANA----- 454
Db 308 KKLGYDGNLLAEYORGTTIDASTYTAINTALAIAGVSAANAAGSVIASPLALLVSGIT 367
QY 455 ----- 454
Db 368 GYISTIILOYSKQAMFEHVANKIHNKIYWEKNNHGNKFENGYDARYLANLODNKKFLN 427
QY 455 ----- 454
Db 428 LNKELQARVIAITQQQMDNNIGLAGISRLGEKVLSCAKAYVDAFEEGKHITKADKLVOLD 487
QY 455 ----- 454
Db 488 SANGIIDVSNAGKAKTOHILFPTPLPGTEHREVRVOTGKYEYITKLNINRDSWKITDG 547
QY 455 ----- 454
Db 548 AASTFDLTNNVQRIELDNAGAVTKETKIITAKLGEDDNVFVSGTTEIDGEGEYD 607
QY 455 ----- 454
Db 608 RVHYSRGNYGALTIDATKETEFGSYTVNRFVETGKALHEVSTHPTALGNREKIEYRHS 667
QY 455 ----- 454
Db 668 NNQHHGYTDTLKAVEIIGTSHNDIFKSGKFNDAENGDDVDTIDGNDGNDRLFQKG 727
QY 455 ----- 454
Db 728 GDDILDGNGDDFDLGGKNDLHGKGGDIFVHRKGGNDIITDSDGNDKLSFSDSNLK 787
QY 455 DLTEFKVHNLYVITNSKKEKVTIQMWFREADFAKEVPYKATKDEKIEIIGONGERITS 514
Db 788 DLTEFKVHNLYVITNSKKEKVTIQMWFREADFAKEVPYKATKDEKIEIIGONGERITS 847
QY 515 KOVDDLAKGNGKITODELSKVVDNYELKHSKNVTNSLDKLSSVSFAFTSSNDSRNLY 574
Db 848 KOVDDLAKGNGKITODELSKVVDNYELKHSKNVTNSLDKLSSVSFAFTSSNDSRNLY 907
QY 575 APTSMLDQSLSSLOFARGSQ 594
Db 908 APTSMLDQSLSSLOFARGSQ 927

RESULT 15
US-07-908-253-2
; Sequence 2, Application US/07908253
; Patent No. 5534256
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: HARLAND, RICHARD J.
; TITLE OF INVENTION: HAEMOPHILUS SOMNUS OUTER MEMBRANE
; TITLE OF INVENTION: PROTEIN EXTRACT ENRICHED WITH IRON-REGULATED PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROBERTA L. ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO

```

STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/908,253
FILING DATE: 19920702
CLASSIFICATION: 420
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9000-0026
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 926 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-908-253-2

Query Match 59.2%; Score 2126; DB 1; Length 926;
Best Local Similarity 52.6%; Pred. No. 2.1e-152;
Matches 483; Conservative 0; Mismatches 0; Indels 436; Gaps 1;

OY 111 SFPKTSKAKKIIYIPONYQYDEQNGLODLYKAAEELGIEYREERNINATQTSIGTI 170
|||||
DB 8 SFPKTSKAKKIIYIPONYQYDEQNGLODLYKAAEELGIEYREERNINATQTSIGTI 67
|||||
OY 171 QTAIGLTERGIIVASAPQIDKLQKTAKGALGSAESIIVONANKAKTVLSGIGSVLA 230
|||||
DB 68 QTAIGLTERGIIVASAPQIDKLQKTAKGALGSAESIIVONANKAKTVLSGIGSVLA 127
|||||
OY 231 GMDLDELQNSNQHALARAGLELTNSLTENINSVKTIIDEPGEQISQPSKIQNTKGLG 290
|||||
DB 128 GMDLDELQNSNQHALARAGLELTNSLTENINSVKTIIDEPGEQISQPSKIQNTKGLG 187
|||||
OY 291 TLGDKLKNIGGLDAGLDIVISGLSGATAALVLDKNASTAKKVGAGFELANQVYNI 350
|||||
DB 188 TLGDKLKNIGGLDAGLDIVISGLSGATAALVLDKNASTAKKVGAGFELANQVYNI 247
|||||
OY 351 TKAVSSYIIAQRYAAGLSSTGPAALIASTVSLAISPLAFAGIADKFNHAKSLESTAE 410
|||||
DB 248 TKAVSSYIIAQRYAAGLSSTGPAALIASTVSLAISPLAFAGIADKFNHAKSLESTAE 307
|||||
OY 411 KRLGYGDNLLEYORGTGTIDASYTAINTALAAIAGVSAANA----- 454
|||||
DB 308 KRLGYGDNLLEYORGTGTIDASYTAINTALAAIAGVSAANA----- 367
|||||
OY 455 ----- 454
DB 368 GVIETILQYSKQAMFEHVANKINKIVEMEKNNHKNYPENGDAYRLANLQDNMKFLLN 427

OY 455 ----- 454
DB 428 LNKELQAEVIAITQOQWNNIGDLAGISRLGEKVLGKAYVDAFEKGHIKADKLIVOLD 487

OY 455 ----- 454
DB 488 SANGIIDVNSGAKAQOHLIFETPLTPGTEHREVRVOTGKEYEITKLNINRVDSMKITDG 547

OY 455 ----- 454
DB 548 AASSTFDLINVVQRIELEDNAGNVTKTETKIIIAKLGEGDDNVFVGSCTTEIDGGEYD 607

OY 455 ----- 454

DB 608 RVHSRNGALTIIDATKETEQGSYTVNRFVETGKALHEVYSTHALVGNREEKIEYRHS 667
455 ----- 454
DB 668 NNOHAGYTTKDTLKAVEEIIIGTSHNDIFKGSKEFNDAPNGGCVDTIIDGNDRLFGK 727
455 ----- 454
DB 728 GDDIIDGNGDDFIDGKGNDLHAGKGDDIFVHRKGDGNDIITDSGNDKLSFSDSNLK 787
OY 455 DLTFEYKHNLVITNSKKEKVTIQQNFREADFAKEVPNPKATKDEKIEEIIIGNGERITS 514
|||||
DB 788 DLTFEYKHNLVITNSKKEKVTIQQNFREADFAKEVPNPKATKDEKIEEIIIGNGERITS 847
OY 515 KOVDDLIAKNGKITQDELISKVVDNYELKHSKNVTNSLDKLISVSAPFSSNDSRNLV 574
|||||
DB 848 KOVDDLIAKNGKITQDELISKVVDNYELKHSKNVTNSLDKLISVSAPFSSNDSRNLV 907
OY 575 APTSMLDQSLSSIQFARGS 593
|||||
DB 908 APTSMLDQSLSSIQFARGS 926

Search completed: March 2, 2001, 10:54:48
Job time: 613 sec

OM of: US-09-306-689-12 to: Issued_Patents_AA.* out_format: pfs
Date: Mar 2, 2001 10:52 AM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet_n2p.model -DEV=xip
-O=/cgn2_1/USPRO.spool/US09306689/runat_02032001_102824_9539/app_query.fasta.1.2389
-DB=Issued_Patents_AA -OPMT=fastan -SUFFIX=rai -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsu62
-TRANS=human40.ctd -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=200000000
-USER=US09306689 -ECGN1_1_98 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:

Query: US-09-306-689-12
Query length: 2088
Database: Issued_Patents_AA.*
Database sequences: 174772
Database length: 17957048
Search time (sec): 62.480000

score_list:

Seq	Strd Orig	ZScore	Escore	Len	Documentation
/cgn2_6/ptodata/2/1aa/5A.COMB	pep:US-08-258-188-2	403.00	574.09	1.7e-25	758
/cgn2_6/ptodata/2/1aa/5A.COMB	pep:US-08-526-813-2	403.00	574.09	1.7e-25	758
/cgn2_6/ptodata/2/1aa/5A.COMB	pep:US-07-690-983D-47	368.00	545.39	5.9e-23	84
/cgn2_6/ptodata/2/1aa/5A.COMB	pep:US-08-694-865-4	290.00	435.73	1.3e-16	49
/cgn2_6/ptodata/2/1aa/5B.COMB	pep:US-08-878-748-4	290.00	435.73	1.3e-16	49
/cgn2_6/ptodata/2/1aa/5B.COMB	pep:US-08-878-748-4	290.00	435.73	1.3e-16	49

seq_documentation_block:

Sequence 2, Application US/08258188
Patent No. 5475098
GENERAL INFORMATION:
APPLICANT: HALL, Robert H.
APPLICANT: XU, Jian Guo
TITLE OF INVENTION: A NEW AND DISTINCTIVE DNA SEQUENCE OF E.
TITLE OF INVENTION: coli 0157:H7 AND ITS USE FOR THE RAPID, SENSITIVE AND
SPECIFIC DETECTION OF 0157:H7 AND OTHER ENTEROHEMORRHAGIC
TITLE OF INVENTION: E. coli
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/258,188
FILING DATE: 14-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Garrett-Wackowski, Eugenia
REGISTRATION NUMBER: 37,330

REFERENCE/DOCKET NUMBER: 15280206, DHSSEL35940
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 543-9600

TELEFAX: (415) 543-5043

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 758 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-258-188-2

alignment_scores:

Quality	Ratio	Length	Gaps
403.00	1.807	395	11
Percent Similarity: 56.456	Percent Identity: 30.633		

alignment_block:

US-09-306-689-12 x US-08-258-188-2

Align seg 1/1 to: US-08-258-188-2 from: 1 to: 758

964	GCACCTGTAAGTCCAGATATAAATGCTTCAACAGCTAAAGTGGTGC	1013
1	SerPheIleuclYlnsSerSapAlaHlStHrGlyThrlYlsAlaIal	17
1014	GGGTTTGAATTGGCAACCAAGCTTGGTAATATACCAAGCCGTT	1063
17	aglyIleGlulLeuThrThrGlValIleuGlYlnsValIglYlsAlaIals	34
1064	CTCTTACATTTAGCCCAACGCTTGACAGCGTTATCTTAACGTGG	1113
34	erGlnTrIleuAlaIalArgMetAlaGlnIleuSerThrAla	50
1114	CCGTGGCTGCTTAAATGCTTCTTACGTTTCTTGCAGTTAGCCATT	1163
51	AlaSerAlaGlyleuIleHrSerAlaValMetLeuAlaIleSerProle	67
1164	AGCAATTTGCCGATATTCGCGATTAATTAATCAATGCAAAAGTTAGAGA	1213
67	userPheLeuAlaIalAspIysPheGluArgAlaIalsGlnLeuGln	84
1214	GTTATGCCGACGCTTTAAATAATTAGCGTATGACGAGTAATTTATTA	1263
84	erTrSerGlnArgPheLysLysLeuAsnTrIgluIylAspAlaLeuLeu	100
1264	GCAGATATCAGCGGGAACAGGACTATTGATCGGTTACTGTCGCAAT	1313
101	AlaAlaPheHisLysGluThrGlyAlaIleAspAlaLeuThrTrI	117
1314	TATACCGCATGGCCGCTATTCGCTGCTGCTGCTGCTGCTGACCC	1362
117	eaSnThrValIleuSerSerValSerAlaGlyIalSerAlaIalSerSera	134
1362	1362
134	laserLeuIleGlyAlaProIleSerMetLeuValSerAlaLeuThrGly	150
1363GATTAACTTTGAAAGT	1382
151	ThrIleSerGlyIleLeuGluAlaSerLysGlnAlaMetPheGluHisVa	167
1383	TAAACATAATCTTGTCATCAGCAAGTAAAGAAAGAAAGTGCACCATTC	1432
167	lAlaGluLysPheAlaIalArgIleAsnGlnTrIgluLysGlnHisGlyL	184
1433	AAACGTGTTCCGAGAG	1449
184	YasnTrPheGluAsnGlyTrYAspAlaArgHisAlaIalPheLeuGlu	200
1450GCTGATTTTGCTAAAGAGAGTCCGTAATATATA	1481

```

201 AspSerLeuLeuAlaAspPheSerArgGlnHisAlaValGluAr 217
1482 AGCAACTAATA.....GATGAGAAATCGAAGAAATCA 1513
      |||:|||||:|||||:|||||:|||||:|||||:
217 gAlaValAlaIleThrGlnGlnHisTrpAspGluLysIleGlyIuLeuA 234
1514 TCGGT.....CAAAATGGCGAGCGGATCACCTCAAGCAAGTTGAT 1557
      |||:|||||:|||||:|||||:|||||:|||||:
234 laGlyIleThrArgAsnAlaAspArgSerGlnSer..... 245
1558 CTATTCGCAAAAGGTAAAGCGCAAAATTAACCAAGATGACATATAAGT 1607
      |||:|||||:|||||:|||||:|||||:|||||:
246 .....GlyLysAlaTyrIle.....AsnTyr 252
1608 TGTGTATACATATGATTCGTCAAA...CATAGCAAAATGTGACAACA 1654
      |||:|||||:|||||:|||||:|||||:|||||:
252 rLeuGlnAsnGlyIleLeuLeuGlnAlaGlnProLysGlnPheThrGln 269
1655 GCTTA...GATTAAGTAACTCATCTGTAAAGTCATTACCTGCTTAAT 1701
      |||:|||||:|||||:|||||:|||||:|||||:
269 InValPheAspProGlnLysGlyThrIleAspLeuSerThrGlyAsnVal 285
1702 GATTGCAAAATGATTAATAGTGGCTCAACTTCATATGTTGATCAAAAGTTT 1751
      |||:|||||:|||||:|||||:|||||:|||||:
286 SerSerValIleThrPheIleThrProThr..... 295
1752 ATCTTCTCTCAATTTGCTAGGGGATCTCAAGCATTTGAGCTAGCGCTGC 1801
      |||:|||||:|||||:|||||:|||||:|||||:
296 .....PheT 297
1802 GCCCTGGCAGCGGTTCTCAAGATTTGAGCTAGCGCTGCCTCGGTCGC 1851
      |||:|||||:|||||:|||||:|||||:|||||:
297 hrProGlnGluValArgGlu..... 304
1852 TCTAGCCAGCATTTGAGCTAGCGCTGCCTCGGTCAGCGCTAGCAGA 1901
      |||:|||||:|||||:|||||:|||||:|||||:
305 .....ArgLysGlnSerGlyTyrG1 312
1902 TTGAGAGTACGCGCTGGCTCGGCTGATGATTCAGCATTTGAGCTAGC...G 1948
      |||:|||||:|||||:|||||:|||||:|||||:
312 uTyMetThrSerLeuIleValAsnGlyLysAspThrTrpSerValLysG 329
1949 GCGTCGCGCTGGCAGCGGTTCTCAAGATTTGAGC 1983
      |||:|||||:|||||:|||||:|||||:|||||:
329 LyIleLysAsnHisLysGlyValLysAspTyrSer 340
seq_name: /cgn2.6/ptodata/2/1aa/5A_COMP.pcp:US-08-526-813-2
seq_documentation_block:
; Sequence 2, Application US/08526813
; Patent No. 5756293
; GENERAL INFORMATION:
; APPLICANT: Hall, Robert H.
; APPLICANT: Xu, Jian Guo
; TITLE OF INVENTION: A New and Distinctive DNA Sequence of E.
; Patent No. 5756293
; TITLE OF INVENTION: coli 0157:H7 and its Use for the Rapid, Sensitive and
; TITLE OF INVENTION: Specific Detection of 0157:H7 and Other Enterohemorrhagic
; TITLE OF INVENTION: E. coli
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/526,813
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,188
; FILING DATE: 14-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 15280-206-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 758 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-526-813-2

alignment_scores:
      quality: 403.00      length: 395
      ratio: 1.807      gaps: 11
Percent Similarity: 56.456      Percent Identity: 30.633

alignment_block:
US-09-306-689-12 x US-08-526-813-2 ..

Align seq 1/1 to: US-08-526-813-2 from: 1 to: 758

964 GCACTGTACTTGCGATTAATAATGCTTCAACAGCTAAAGTGGGTGC 1013
1 SerPheIleLeuGlyAsnSerAspAlaHisThrGlyThrLysAlaIleAl 17
1014 GGGTTTGAATTGGCAACCAAGTTGTGTAAATATTACCAAGCCGTTT 1063
17 agLyIleGlnIleuThrThrGlnValIleuGlnAsnValGlyLysAlaValS 34
1064 CTCTTTCATTTTACCCCAACGTTTGCACAGGTTTATCTTCAACGCG 1113
34 erGlnTyrIleLeuAlaGlnArgMetAlaGlnGlyLeuSerThrThrAla 50
1114 CCTGGCTGCTTAAATGCTTACTGTTCTCTTCCGATTAAGCCATT 1163
51 AlaSerAlaGlyLeuIleThrSerAlaValMetLeuAlaIleSerProle 67
1164 AGCAATTCGCGGTAATTCGCGATAATTTAATCATGCAAAAAGTTAGAGA 1213
67 uSerPheLeuAlaIleAlaAspLysPheGluArgAlaLysGlnLeuGlnS 84
1214 GTTATGCCGAGCGTTTAAATAATTTAGCTATGACGAGATTAATTATA 1263
84 erTyrSerGlnArgPheLysLysLeuAsnTyrGlnGlyAspAlaLeuLeu 100
1264 GCAAAATATAGCGCGGAGACAGCATTTGATGATCGGTTACTGCAAT 1313
101 AlaIaPheHisLysGlnThrGlyAlaIleAspAlaIaIeuthrThrI1 117
1314 TAATACCGCATTTGCGCGCTATTCGTGTGTGTCTGCTGCAGCC. 1362
117 eaSnThrValLeuSerSerValSerAlaGlyValSerAlaIaIeSerSera 134
1362 ..... 1362
134 laSerIleuIleGlyAlaProIleSerMetLeuValSerAlaLeuThrGly 150
1363 .....GATTAACTTTGAAAAGT 1382
151 ThrIleSerGlyIleLeuGlnAlaSerLysGlnAlaMetPheGlnHisVa 167
1383 TAAACATATCTTGTCATCAAGATAGCAAAAAGAGAAAGTGCATTC 1432

```

```

167 lalagluyspheaalaalaargileasnclutrrgluylusglunlslgl 184
1433 AAAATGCTCCGAGAG..... 1449
184 yAAsnltYrhegluAsnclYtYrAspAlaRghIsAlaAlaRheleucln 200
1450 ..... GCTGATTTTGGTAAAGAGTGCCTAATTATPA 1481
201 AspSerleuSerleuLeuAlaAspRheSerArglnHisAlaValclut 217
1482 AGCAACTAAA..... GATGAGAAATTCGAAGAAATCA 1513
217 galavalAlaIlethrglnlnlnlstrpsrpluylslleglgluleva 234
1514 TGGGT.....CAAAATGGCGAGCGGATCACTCAAGCAAGTGTGAT 1557
234 lAglylYlethrrArgAsnAlaAspArgSerGlnSer..... 245
1558 CTTATCGCAAAAGGTAAAGCAAAATTACCAAGATGAGCTATCAAAAGT 1607
246 ..... GlylYsAlaTyrIle.....AsnTyr 252
1608 TCGTATATATGATGATTCCTCAAA...CATAGCAAAATGTCACAACA 1654
252 rleugluAsnclYglYleuLeuclnAlaInProlysgluRhetHrgln 269
1655 GCTTA...GATAAGTTAATTCATCATCTGTAAGTCATTAACCTGCTTAAT 1701
269 lvalRheAspProclnYsglyThrIleAspLeuSerThrglyAsnVal 285
1702 GATTGAGAAATGTATAGTGGTCCAACTTCATGTTGGATCAAAAGTTT 1751
286 SerSerValleuThrRheIleThrProthr..... 295
1752 ATCTTCTTTCATTTGCTAGAGGATCTCAGCATTCGAGCTACGGCTGC 1801
296 ..... Phet 297
1802 GCCCTGGCAGCGGTTCTCAAGATTGGAGCTACGGCTCGCGGGTGC 1851
297 hPrrglYglclnValArgIlu..... 304
1852 TCTAGCAGCATTTGGAGCTACGGCTCGCGGCGAGCGGATGAGCAAGA 1901
305 ..... ArgLYsGlnSerGlyLysTyrcl 312
1902 TTGGAGCTACGGCTCGCTCCGGTGGATCTCAGCATTCGAGCTAC...G 1948
312 uYrmetHrSerleuIleValAsnclYlYsAspHrrTrpSerValYsg 329
1949 GCCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAGC 1983
329 lYlYleYsAsnHlslYsglyValYlYrAspTyrSer 340
seq_name: /cgn2_6/ptodata/2/1aa/PCBUS_COMB.pep: PCT-US95-08554-2
seq_documentation_block:
; Sequence 2, Application PC/TUS9508554
; GENERAL INFORMATION:
; APPLICANT: HALL, Robert H.
; APPLICANT: XU, Jian Guo
; TITLE OF INVENTION: A NEW AND DISTINCTIVE DNA SEQUENCE OF E.
; TITLE OF INVENTION: coli 0157:H7 AND ITS USE FOR THE RAPID, SENSITIVE AND
; TITLE OF INVENTION: SPECIFIC DETECTION OF 0157:H7 AND OTHER ENTEROHEMORRHAGIC
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US

```

```

? ZIP: 94105-1493
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentln Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US95/08554
? FILING DATE: 14-JUN-1994
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: Garrett-Wackowski, Eugenia
? REGISTRATION NUMBER: 37,330
? REFERENCE/DOCKET NUMBER: 15280206, DHHS135940
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 543-9600
? TELEFAX: (415) 543-5043
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 758 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? PCT-US95-08554-2

alignment_scores:
Quality: 403.00 Length: 395
Ratio: 1.807 Gaps: 11
Percent Similarity: 56.456 Percent Identity: 30.633

alignment_block:
US-09-306-689-12 x PCT-US95-08554-2 ..
Align seg 1/1 to: PCT-US95-08554-2 from: 1 to: 758

964 GCACCTGTACTTGCAGATAAAGCTTCACAGCTAAAGAGTGGGTGC 1013
1 SerPheIleuGlyAsnSerAspAlaHlstrnglyThrLYsAlaAlaI 17
1014 GGGTTTGAATTGGCAACCAAGTTGTTGTAATATTACCAAGCCGTT 1063
17 aglylYleuThrThrglnValleuGlyAsnValglYlYsAlaValS 34
1064 CTTCTTACATTTAGCCCAAGCTGTTCAGCAGGATTATCTTCACTGG 1113
34 erGlnIYrIleleuAlaglnArgmetAlaglnclYleuSerThrrAla 50
1114 CCGTGGCTGCTTAATGCTTCTACTGTTCTGCTTGGCATAGCCCAT 1163
51 AlAserAlaglyleuIleThrSerAlaValMetLeuAlaIleSerProle 67
1164 AGCATTTGCCGATATTGCCGATTAATTATCATCATCAAGAAAGTTAGAGA 1213
67 userPheleuAlaAlaAlaAspLysPheclnArgAlaLYsGlnleuGlnS 84
1214 GTTAGCCGACGCTTAAATAATTAGCTATGACGAGCATATAATTATTA 1263
84 ertYserGlnArgPheLysLysLeuAsnTyrclnclYlYsAlaLeu 100
1264 GCAGATATACGGGGGAGGAGGAGTATGATGATGATGCTGATACGCAT 1313
101 AlAlaAlaPheHlslYsglyThrGlyAlaIleAspAlaAlaLeuThrI 117
1314 TATATCCGATTTGCCGCTATTGCTGCTGCTGCTGCTGCTGACGC 1362
117 easnThrValleuSerSerValSerAlaIleSerAlaIleSerSerA 134
1362 ..... 1362
134 lAserleuIleGlyAlaProIleSerMetLeuValSerAlaLeuThrgly 150
1363 ..... GATTTAACATTGAAAAAGT 1382

```

```

151 ThrIleSerGlyIleuGluAlaSerLysGlnAlaMetPheGluHisVala 167
1383 TAACACATATCTGTGTCACAGATAGCAAAAAGAGAAAGTGCACATTC 1432
167 LAlaGluLysPheHisAlaAlaArgIleAsnGluTrpGluLysGlnHisGlyL 184
1433 AAAAGTGTTCGAGAG..... 1449
184 ysaSnTrpPheGluAsnGlyTrpAspAlaArgHisAlaAlaPheLeuGlu 200
1450 .....GCTGATTTCCTAAGAACTGCTAATTATTA 1461
201 AspSerLeuSerLeuAlaAspPheSerArgGlnHisAlaValGluArg 217
1482 AGCAACTAAA.....GATGAGAAATGCAAGAAATCA 1513
217 gAlaValAlaIleThrGlnGlnHisTrpAspGluLysIleGlyGluLeuA 234
1514 TCGGT.....CAAAATGGCGAGCGATCACCTCAAGCAAGTATGAT 1557
234 laGlyIleThrArgAsnAlaAspArgSerGlnSer..... 245
1558 CTATTCGCAAAAGGTACGCGCAAAATACCAAGATGATGCTATCAAAAGT 1607
246 .....GlyLysAlaTrpIle.....AsnTrp 252
1608 TGTTGATTAATGATGATGCTCAAA.....CATAGCAAAATGTGCAAAACA 1654
252 rLeuGlnAsnGlyGlyLeuLeuGlnAlaGlnProLysGluPheThrGlnG 269
1655 GCTTA...GATTAAGTTAATCTCATCTGTAAGTGCATTACCTGCTAAT 1701
269 lValAlaPheAsnProGlnLysGlyThrIleAsnLeuSerTrnGlyAsnVal 285
1702 GATTCGAGAAATGATTAGTGGCTCCAACTTCATGTTGATCAAAAGTTT 1751
286 SerSerValLeuThrPheIleTrnProThr..... 295
1752 ATCTTCTCTCAATTTGCTAGGGGATCTCAGCAATTTGAGCTAGCGCCTGC 1801
296 .....PheT 297
1802 GCCCTGGCAGCGGTTCTCAAGATTGAGCTAGCGGCTGCGGTTGCG 1851
297 hrProGlyGlnGluValArgLys..... 304
1852 TCTAGCCAGATTGGAGCTAGCGGCTGCGGCTGGCAGCGTACCAAGA 1901
305 .....ArgLysGlnSerGlyLysTrpGly 312
1902 TTGGAGCTAGCGGCTGCGGCTGCGGCTGATGATGAGCTAGGAGCTAC...G 1948
312 uTrpMetThrSerLeuIleValAsnGlyLysAspThrTrpSerValLysG 329
1949 GCCTGCGGCTGGCAGCGGTTCTCAAGATTGAGAC 1983
329 LyIleLysAsnHisLysGlyValTrpArgTrpSer 340
seq_name: /cgn2_6/plodata/2/1aa/5A_COMB.pep:US-07-690-983D-47
seq_documentation_block:
; Sequence 47, Application US/07690983D
; Patent No. 5403586
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, Gregory J.
; APPLICANT: STEWART, Andrew G.
; APPLICANT: TSONIS, Con G.
; TITLE OF INVENTION: FUSION PROTEINS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: F. J. Lardner
; STREET: 3000 K Street, N.W.

```

```

; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/690,983D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00373
; FILING DATE: 24-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/148 CHAC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-690-983D-47

```

```

alignment_scores:
  Quality: 368.00      Length: 101
  Ratio: 4.779        Gaps: 7
  Percent Similarity: 76.238  Percent Identity: 68.317

```

```

alignment_block:
US-09-306-689-12 x US-07-690-983D-47 ..

```

```

Align seg 1/1 to: US-07-690-983D-47 from: 1 to: 84

```

```

22 TCTCAGCATTTGAGACTACGGCTGCGCTGGCAGCGGTTCTCAAGTTG 71
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2 SerGlnHisTrpSerTrpGlyLeuArgProGly.....GlnHisTr 15
72 GAGCTAGCGGCTGCGGCTGCGGCTGCTAGCAGCATTTGAGCTAGCGGC 121
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
15 pSerTrpGlyLeuArgProGly.....GlnHisTrpSerTrpGlyL 29
122 TGCAGCTGCGCAGCGGTTAGCCAAAGATTGAGCTAGCGGCTGCGGGT 171
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
29 euArgProGly.....GlnHisTrpSerTrpGlyLeuArgProGly 42
172 GAGCTCAGCATTTGAGACTACGGCTGCGGCTGGCAGCGGTTCTCAAGA 221
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
43 .....GlnHisTrpSerTrpGlyLeuArgProGly.....GlnH 54
222 TTGGAGCTAGCGGCTGCGGCTGCGGCTGCTAGCAGCATTTGAGCTAGC 271
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
54 sTrpSerTrpGlyLeuArgProGly.....GlnHisTrpSerTrpG 68
272 GCCTGCGGCTGGCAGCGGTTAGCCAAAGATTGAGCTAGCGGCTGCGGC 321
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
68 LyLeuArgProGly.....GlnHisTrpSerTrpGlyLeuArgPro 81
322 GGT 324
   |||
82 Gly 82

```

```

seq_name: /cgn2_6/plodata/2/1aa/5A_COMB.pep:US-08-387-156-4
seq_documentation_block:
; Sequence 4, Application US/08387156

```

```
; Patent No. 5723129
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HOW P.A.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/387,156
; FILING DATE: 10-FEB-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/960,932
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,171
; FILING DATE: 16-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 9001-0016.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 617-8999
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-387-156-4

alignment_scores:
  Quality: 290.00      Length: 49
  Ratio: 5.918        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-306-689-12 x US-08-387-156-4  ..
Align seg 1/1 to: US-08-387-156-4 from: 1 to: 49

25 CACGATTGAGCTACGCGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 74
|||||
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 17
rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34

75 CTACGCGCTCGTCGCGGTGGTCTTACCCAGCATTTGAGCTACGCGCTGC 124
|||||
17 rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34

125 GCCCTGGCAGCGGTAGCCAAAGATTGGAGCTACGCGCTCGCGGT 171
|||||
34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-694-865-4
seq_documentation_block:
; Sequence 4, Application US/08694865
; Patent No. #5837268
; GENERAL INFORMATION:
```

```
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,865
; FILING DATE: 09-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-694-865-4

alignment_scores:
  Quality: 290.00      Length: 49
  Ratio: 5.918        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-306-689-12 x US-08-694-865-4  ..
Align seg 1/1 to: US-08-694-865-4 from: 1 to: 49

25 CACGATTGAGCTACGCGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 74
|||||
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 17
rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34

75 CTACGCGCTCGTCGCGGTGGTCTTACCCAGCATTTGAGCTACGCGCTGC 124
|||||
17 rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34

125 GCCCTGGCAGCGGTAGCCAAAGATTGGAGCTACGCGCTCGCGGT 171
|||||
34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-878-748-4
seq_documentation_block:
; Sequence 4, Application US/08878748
; Patent No. 5969126
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: REDMOND, MARK J.
; APPLICANT: HUGHES, HOW P.A.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 635 BRYANT STREET
; CITY: PALO ALTO
```

```

STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,748
FILING DATE: 19-JUN-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-748-4

```

```

alignment_scores:
Quality: 290.00      Length: 49
Ratio: 5.918        Gaps: 0
Percent Similarity: 100.000    Percent Identity: 100.000

```

alignment_block:
US-09-306-689-12 x US-08-878-748-4 ..

Align seg 1/1 to: US-08-878-748-4 from: 1 to: 49

```

25 CAGATGAGCTACGCGCTGCGCTGCTCAAGATTGAG 74
|||||
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnAspTrpSe 17
75 CTACGCGCTGCGCTGCGCTGCTGAGCAGATTGAGCTACGCGCTGC 124
|||||
17 rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34.
125 GCCCTGCGAGCGGTAGCCAGATTGAGCTACGCGCTGCGGCT 171
|||||
34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49

```


[illegible]

RESULT	2
09LC58	
ID	09LC58
AC	09LC58;
DT	01-OCT-2000 (TREMBLrel. 15, Created)
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE	HEMOLYSIN A.
CN	EHCC-HLTA.
OS	Escherichia coli.
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC	Escherichia.
OX	NCBI_TaxID=562;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Tanabe I., Wakisaka-Saito N., Harada Y., Zhang H.-M., Yamamoto T.;
RT	The enterohemorrhagic Escherichia coli (EHEC)-hemolysin genes of a
RT	Shiga toxin 1 (Stx1)- and Stx2-producing, serotype O128 Escherichia
RT	coli strain with a greatest hemolytic activity.";
RL	Acta Med. Biol. 0:0-0(2000).
SR	EMBL: AB032930; BAA93708.1; -
SO	SEQUENCE 998 AA; 107196 MW; CD7A8BE9BD62DB6 CRC64;

Query Match	23.8%;	Score 855;	DB 2;	Length 998;
Best Local Similarity	35.8%;	Pred. NO. 1e-36;		
Matches 220;	Conservative 92;	Mismatches 182;	Indels 120;	Gaps 15;

[illegible]

Db	262	TTQYLGANVGRASQYLLAQBRAQSLSTTASAGLITSAVWLALSPLSFLIADKFEBAKQ	321
QY	403	LESTAERERKKLGYGDMLLAEYORGTSTIDASVTAIN TALAIAGVSAANA-----	454
Db	322	LESYERERKKLNTYEGDALLAAFFHEKETAIDALETITNTVLSAGVSAASSASLIGAPL	381
QY	455	-----DLTFPKYVHNLYITNSKKEKXTIONPRE-----	483
Db	382	SMLVSAITGTTISGILLEASKOAMFEHVADKPAARINEMKEHKNYFENGVDARHAAFL	441
QY	484	-----ADFAKEVPWYKATK-----DEKIEEIG--ONGERITSKVDLIDIAKNGKITO	530
Db	442	SLSLADFSRKHANERAVAITQOHMDEKIGELAGITFNADRSQS-----GKAYI--	490
QY	531	DELSKYVDNVELLK-HSKNVTNSL-DKLISSVSAFTSSNDNRVLVAPTSMIDQSLSLQ	588
Db	491	-----NYLENGLLBAQKREFIQVFDPQKGTIDLSTGVNWSVLTFLVPT-----	535
QY	589	FARQSQHMWSTGLRPGSSQDMSTVGLRPGSSQHMWSTVGLRPGSSQDMSTVGLRPGSSQHM	648
Db	536	-----FTPGGEVRE-----RKOSKYEYEMTSLVINGKDTWS	566
QY	649	X-GLRPGSGSQDWS	661
Db	567	VKGITKNHKGVIYDYS	580

RESULT	3			
085101				
ID	085101	PRELIMINARY;	PRT;	998 AA.
AC	085101;			
DT	01-NOV-1998 (TREMBLrel. 08, Created)			
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)			
DT	01-JUN-2000 (TREMBLrel. 14, Last annotation update)			
DE	HEMOLYSIN.			
GN	EHXA.			
OS	Escherichia coli.			
OC	Plasmid EHEC-hemolysin.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Escherichia.			
OX	NCBI_TaxID=562;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ENTEROHEMORRHAGIC EC920006;			
RX	MEDLINE=98261495; Pubmed=9596716;			
RA	Boerlin P., Chen S., Colbourne J.K., Johnson R., De Grandis S.,			
RA	Gyles C.;			
RT	"Evolution of enterohemorrhagic Escherichia coli hemolysin plasmids			
RT	and the locus for enterocyte effacement in shiga toxin-producing E.			
RT	coli.";			
RL	Infect. Immun. 66:2553-2561(1998).			
RL	EMBL; AF043471; AAC24352.1; -.			
DR	INTERPRO; IPR001343; -.			
DR	PFAM; PF00353; hemolysincabind; 2.			
DR	PRINTS; PR00313; CABNDNGRPT.			
DR	PROSITE; PS00330; HEMOLYSIN_CALCIIUM; 2.			
KW	plasmid.			
QC	SEQUENCE			
	998 AA; 107095 MW; ESA566B239DCBE55 CRC64;			

Query Match	23.6%	Score 849;	DB 2;	Length 998;
Best Local Similarity	35.7%	Pred. No. 2.1e-36;		
Matches 219; Conservative	92;	Mismatches 183;	Indels 120;	Gaps 15;

	Query Match	23.6%	Score 849	DB 2:	Length 998
	Best Local Similarity	35.7%	Pred. No. 2.1e-36		
	Matches 219	Conservative 92	Mismatches 183	Indels 120	Gaps
Qy	110	SSPFTGAKRIITYIPONYDYDFGCGGLDDLVAAAEELGIEVQREBRNNIATAPQSLGT	169		
		: : : : : : : : : : : : : : :			
Db	25	SSSVRSKSGKRIILILIPNTE---AQGVGIELVKAADDELGIETHRERDDTALIANPFET	81		
		: : : : : : : : : : : : : : :			
Qy	170	IQTATIGLTERGVILSAPOIDKLQK--TKAAQAL--GSAESTVQANAKRAFTLSGISOILGS	227		
		: : : : : : : : : : : : : : :			
Db	82	AEKVVGLTERGVALFAAQDQLKLLQKQKVGSKIGGTAEVNNGNLAKRAGVYLSLAKQNFGTG	141		
		: : : : : : : : : : : : : : :			
Qy	228	VLAGMDIDELAQ-----NNSNQHLAKAKAGLETNLSLENIANSVKTLDDEFGEOISQFSK	282		
		: : : : : : : : : : : : : : :			

```

Db 142 ALSGMALDELLRKOREGDISQNDIAKSSIELINQLVDTVSSINSTVDSFSQLNQLCSF 201
Qy 283 LONIKGLCTLDKLNIGGLDCKAGLDVIGSLGSAFALVADKNASTAKVAGFEL 342
Db 202 LSKRRLSSVGKQLNLDPLGRLGDLVVGSLGSAVSFLGNSDHTGKAAGIEL 261
Qy 343 ANOVGNITKAVSSYLLAQRVAAAGLSTGPVAAALIASTVLAISPLAFAGIADKFNHKS 402
Db 262 TTYVAGNVKAVSQYLLAQRMAQGLSTPAASAGLITSAVMLAISPFLAIDKFERAKQ 321
Qy 403 LESYAEERKKTIGYDGNLLAEYORGTITDASVTAINALAIAGVAAAAA----- 454
Db 322 LESYERKKTINVEDALLAFHRESGALDALTITNVLSVSGVSAASASLIGAPI 381
Qy 455 -----DLTFEKVKHNLVITNSKKEKVTIOWFRE----- 483
Db 382 SMLVSALTGTISGLEASKQAMFEHVADKFAARINEMEKKNYFENGVDARHAAFLSD 441
Qy 484 -----ADFAKEVPNYKATK-----DEKIEELIG--QNGERTISQVODDLAKNGKITQ 530
Db 442 SLSLADFSRQHAAVRAVAITQOHMDEKIGELAGITRRADRQS-----GRAYI-- 490
Qy 531 DELSKVDNVELLK-HSKNVTNSL-DKLISVSATSSNDSRNVLVAPTSMLDQSLSLQ 588
Db 491 -----NYLENGGLLEAQKPEFTQOVFPDPOKGTIDLTSGVSSVLTFITPT----- 535
Qy 589 FARGSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQHWS 648
Db 536 -----FTPEEVEVR-----RKQSGKYEYMTSLIVNGKDTWS 566
Qy 649 Y-GLRPGSSQDMS 661
Db 567 VKGINKHKGVDYS 580

RESULT 4
046716 PRELIMINARY: PRT: 998 AA.
AC 046716;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-MAY-1999 (Tremblrel. 10, last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, last annotation update)
DE HEMOLYSIN A (HLXA).
GN HLXA OR EHEC-HLXA.
OS Escherichia coli.
OC Plasmid p0157.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-0157:H7;
RX MEDLINE=98290540; PubMed=9628576;
RA Makino K., Ishii K., Yasunaga T., Hattori M., Yokoyama K.,
RA Yatsudo H.C., Kubota Y., Yamachi Y., Iida T., Yamamoto K., Honda T.,
RA Han C., Ohtsuda A., Kasamatsu M., Hayashi T., Kohara S., Shinaawa H.;
RT "Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
RT enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
RT outbreak."
RL DNA Res. 5:1-9(1998).
RN [2]
RP SEQUENCE OF 241-998 FROM N.A.
RC STRAIN-EHEC;
RA Hall R.H., Xu J., Walderhaug M.O.;
RL Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-EDL933;
RX MEDLINE=98391744; PubMed=9722640;
RA Burland V., Shao Y., Perna N.T., Plunkett G., Sofia H.J.,
RA Blattner F.R.;
RT "The complete DNA sequence and analysis of the large virulence plasmid

```

```

RT of Escherichia coli O157:H7."
RL Nucleic Acids Res. 26:4196-4204(1998).
DR EMBL; AB011549; BAA31774.1; -.
DR EMBL; U12572; AAA20544.1; -.
DR EMBL; AF074613; AAC70116.1; -.
DR INTERPRO; IPR001343; -.
DR PFAM; PF00353; hemolysinCbind; 2.
DR PRINTS; PR00313; CAMPDNGRPT.
DR PROSITE; PS00330; HEMOLYSIN_CACTUM; 2.
RW Plasmid.
SQ SEQUENCE 998 AA; 107047 MW; BB3C0D36FCBB9EBD CRC64;

Query Match 23.68; Score 847; DB 2; Length 998;
Best Local Similarity 35.58; Pred. No. 2,7e-36;
Matches 218; Conservative 92; Mismatches 184; Indels 120; Gaps 15;

Qy 110 SSFPTGAKKTIIVYPOVYOYDTBOGNGLODVKAELGIEVOREERNIATQTSIGT 169
Db 25 SSSVRSACKKLLILIPNTE--AOGVGINELVRADELGLIEIHTERDDTAIANQFFGA 81
Qy 170 IOTAIGTERGIVLSAPQIDKLLQK-TRAGQAL-GSASEIVQANAKKATVLSGIQSLGS 227
Db 82 AEKVYGLTERGVAIPAPQDLKLLQKQYQKVGKIGSTAEVNGNIGKACTVLSALQNFQI 141
Qy 228 VLAGDDLEALQ-----NNSQHALAKGLEITNSLIANSVKTLDPEGEQISQFSK 282
Db 142 ALSGMALDELLRKOREGDISQNDIAKSSIELINQLVDTVSSINSTVDSFSQLNQLCSF 201
Qy 283 LONIKGLCTLDKLNIGGLDCKAGLDVIGSLGSAFALVADKNASTAKVAGFEL 342
Db 202 LSKRRLSSVGKQLNLDPLGRLGDLVVGSLGSAVSFLGNSDHTGKAAGIEL 261
Qy 403 ANOVGNITKAVSSYLLAQRVAAAGLSTGPVAAALIASTVLAISPLAFAGIADKFNHKS 402
Db 262 TTYVAGNVKAVSQYLLAQRMAQGLSTPAASAGLITSAVMLAISPFLAIDKFERAKQ 321
Qy 484 LESYAEERKKTIGYDGNLLAEYORGTITDASVTAINALAIAGVAAAAA----- 454
Db 322 LESYERKKTINVEDALLAFHRESGALDALTITNVLSVSGVSAASASLIGAPI 381
Qy 455 -----DLTFEKVKHNLVITNSKKEKVTIOWFRE----- 483
Db 382 SMLVSALTGTISGLEASKQAMFEHVADKFAARINEMEKKNYFENGVDARHAAFLSD 441
Qy 484 -----ADFAKEVPNYKATK-----DEKIEELIG--QNGERTISQVODDLAKNGKITQ 530
Db 442 SLSLADFSRQHAAVRAVAITQOHMDEKIGELAGITRRADRQS-----GRAYI-- 490
Qy 531 DELSKVDNVELLK-HSKNVTNSL-DKLISVSATSSNDSRNVLVAPTSMLDQSLSLQ 588
Db 491 -----NYLENGGLLEAQKPEFTQOVFPDPOKGTIDLTSGVSSVLTFITPT----- 535
Qy 589 FARGSQHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPGSSQHWS 648
Db 536 -----FTPEEVEVR-----RKQSGKYEYMTSLIVNGKDTWS 566
Qy 649 Y-GLRPGSSQDMS 661
Db 567 VKGINKHKGVDYS 580

RESULT 5
P71223 PRELIMINARY: PRT: 998 AA.
AC P71223;
DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-FEB-1997 (Tremblrel. 02, last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, last annotation update)
DE EHEC-HEMOLYSIN.
GN EHEC-HLXA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

```

OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=78-92;
RX MEDLINE=97034824; PubMed=880480;
RA Schmidt H., Karch H.;
RT "Enterohemorrhagic phenotypes and genotypes of shiga toxin-producing
RT Escherichia coli O111 strains from patients with diarrhea and
RT hemolytic-uremic syndrome.";
RL J. Clin. Microbiol. 34:2364-2367(1996).
DR EMBL; X94129; CA63849.1; -.
DR INTERPRO; IPR001343; -.
DR PFAM; PF00353; hemolysin_cbind; 2.
DR PRINTS; PR00313; CABINDGPRP.
DR PROSITE; PS00330; HEMOLYSIN_CALCITUM; 2.
SO SEQUENCE 998 AA; 107255 MW; DA3EF078C7E4131E CRC64;

Query Match	23.5%	Score 845	DB 2	Length 998
Best Local Similarity	35.5%	Pred. No.	3,4e-36	
Matches 218	Conservative 92	Mismatches 184	Indels 120	Gaps 15

OY	110	SSPFTGAKKIIITVPOKYVDTEOGNGLODLYAAAEELGEVOREERNRINAFQNSLGT	169
Dd	25	SSSVRSACKKIILLIPDNYE---AQGYINELVKADELGEIETHRTDRDPTALANQFPGA	81
OY	170	IQTAIGLTERGIVSAPOIDKILQK--FKAGALG--SAESIYOVANKAKTVLSGTQSLGS	227
Dd	82	AEKVYGLTERGVAIFAPOLDLQKYQVGSKISRTREBNYGNNGKRGVIALQNPFTGI	141
OY	228	VIAGNDUDEAEO-----NNSMOMLAKRAGELTSSLENTIANSVKTLDEGDEQISOFGSK	282
Dd	142	ALSGADELRLKOREGEDISQNDIASSLELLIQVADTVYSSINSTVDSSEQINDQSGF	201
OY	283	LQNIKGLTGLDKLKNIGDILKAGILGLDIVISGLSLSGATAALVYADKNASTAKVYGAFFEL	342
Dd	202	LSKRRLSSVGKILQNPJDLGSLGDGLDYVSGILSNVSAFPIGNSDAPHGTMAAGIEL	261
OY	343	ANQVGNITTKAVSSYIIAQRVAAGLSTGPYVALLASTVSLAISPLAFACIADKFNHAKS	402
Dd	262	TYOVINGNKAVSOYIIAORMAOGILSTTASAGILTSAVMLAISPLSFLAAADKFERAKO	321
OY	403	LESVAERPKKTYGQDNLLAYEOTGTGIDPASYAINTALAIAGVSAANA-----	454
Dd	322	LESTSERPKKLTNEGDDALLAFKHETGAIDALTTITVYSSVAGVSAHSSASLIGAPI	381
OY	455	-----DLTFEKVKNHLYITNSKEKYTIONMFER-----	483
Dd	382	SNLYSALTLGTISGLIEASKQAMFEHVAEKKFARINEMEKHEGKNYFENGYDARHAAFLD	441
OY	484	-----ADPAKEVPNYKAK-----DEKIEIITG--QNGERITNSKQVDDLIANGKNGITY	530
Dd	442	SLSLSLADESRQHAVERAAVAILTOOHMDKEIGELAGITTTNABRSOS-----GKAVI--	490
OY	531	DELSKVYDNYELK--HSKNVTNSL--DKLISSVSAFTSSSDNRNVLVAPTMLDQSLSSLO	588
Dd	491	-----NYLENGGLLEQKPEFFQYVFDQKGTIDLSIONVSSVLTFFPT-----	535
OY	589	FARQSQHWSYGLRPGSGQSDMSYGLRPGSGQSHWSYGLRPGSGQSDMSYGLRPGSQHWS	648
Dd	536	-----FTPGEEVRE-----RKQSGKEYMFTFLVINGKDTWS	566
OY	649	X-GLRPGSGQSDMS	661
Dd	567	VKGIKNHKGYDYD	580
RESULT	6		
ID	047461	PRELIMINARY;	PRT; 998 AA.
AC	047451;		
DT	01-NOV-1996	(TIMBLREL, 01, Created)	

DT 01-NOV-1996 (TREMBLER, 11, last sequence update)
DE 01-JUN-2000 (TREMBLER, 04, last annotation update)
DE PLASMID-DNA FOR EHEC-HEMOLYSIN OPERON.
GN EHEC-HLYA.
OS *Escherichia coli*.
OC Plasmid p0157.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC *Escherichia*.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EDL 933;
RX MEDLINE=95172699; Pubmed=7868227;
RA Schmidt H., Beutin L., Karch H.;
RT "Molecular analysis of the plasmid-encoded hemolysin of *Escherichia coli* O157:H7 strain EDL 933".
RL Infect. Immun. 63:1055-1061(1995).
DR EMBL; X86087; CAA60042.1; -.
DR INTERPRO: IPR001343;
DR PFAM: PF00353; hemolysinabind; 2.
DR PRINTS: PFO0313; CABNNGRPT.
DR PROSITE: PS00330; HEMOLYSIN_CALCIDIUM; 2.
KW plasmid.
SQ SEQUENCE 998 AA; 107032 MW; 4DBBE108C309BF7E CRC64;

Query Match	23.5%	Score	845	DB 2	Length	998			
Best Local Similarity	35.5%	Pred. No.	3.4e-36						
Matches	218	Conservative	91	Mismatches	185	Indels	120	Gaps	15

QY	110	SSEPTGAKKILIIYPOYQYDTEGNOGLODYAAEELTEIEYERREERNIATQTSIGT	169
Db	25	SSSVASACKKILLILIPNVE---AQGVINELVYAADELGIEIHRTERDOTALANOFPGA	81
QY	170	IQTAIGTERGIVSAPOIDKLLQK-TRAGQAL-GSAESIYONANKAKVYLSGTSIILGS	227
Db	82	AEKVGLTERGVAIFAPOLDKLLOKYOKVSGSIGTGEEMNGNLGKAGVIALSALONFTGI	1414
QY	228	VLAGDDIDEALQ-----NNSNMHAAKAGLETSNLIENIANSVKTLDEGGEQISQFGSK	282
Db	142	ALISGALDELILRKOREGEDISQNDIANSIELINQVLDPVWSINSTVDSSEQLNQLGTF	201
QY	283	IONINGLGTLDKLNIGIGDLKAGIGLDVIGSILSGATPAALVLADKNVSAKKGAGFEL	342
Db	202	LSKRRLSSVYGKLONLPLDLPGLDGDVYSGIISVNSASRILGNSDAHGTMAAGIEL	261
QY	343	ANOVGNITKAIVSSYTIIAQRYAAGLSGTGPVAAIIASTVSLAISPLAPAGIADKFHNAKS	402
Db	262	TYQVINGNKAIVSOYIIIAORAAQGSITSTAASAGLITSAMVLAIISPISFLAADKFERAKO	321
QY	403	LESYAEERKKIKGYOSDNLAEYOGSTGIDASVAIINPMALAIINGVSAASA	454
Db	322	LESTISERKKLNTBEDALLAFHKEETGALIDGLTTINTVILSSVAGVSAASASLIGAPI	801
QY	455	-----DLTFEKKVKNLITVTSKKREKVTIONMFE-----	483
Db	382	SMLYSALNCTGISGLEASKQAMFEHVAEKKAFARINEMEKEKKNYFENGYDARHAAFLD	4414
QY	484	-----ADPAKEVPNKATK-----DEKIEBITG--ONGERTITSQYVDLLIAKGNCKITO	530
Db	442	SLSLADFSROHAVERAAVAITHQWDEKITIGELAGITRNADRSOS-----GKAYI--	480
QY	531	DELSKVYDNYNELTK-HSKNNTNSL-DKLISSVAFSTSSNDSRNVLPVAPTSMLDQSSLSLQ	588
Db	431	-----NYLENGLLEQKREFITQYVDPDQKGTIIDSLGTVSSVLFTFPT-----	535
QY	589	FARQSQHSYGLRPGSGSSQSYGLRPGSGSSQHSYGLRPGSGSSQSYGLRPGSGQHS	648
Db	536	-----FTTGEVRE-----RKQSGKEYEMTSLIVGNKDTWS	566
QY	649	Y-GLRPGSGSQDWS	661
Db	567	VKGIRNHKGUYDYS	580

RESULT 7
047262 PRELIMINARY: PRT: 998 AA.
AC 047262;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE HEMOLYSIN.
GN EHEC-HLYA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95172699; PubMed=7868227;
RA Schmidt H., Beutin L., Karch H.;
RT "Molecular analysis of the plasmid-encoded hemolysin of Escherichia coli O157:H7 strain EDL 933."
RL Infect. Immun. 63:1055-1061(1995).
DR EMBL: X79839; CAA56234.1; -;
DR INTERPRO: IPR001343; -;
DR PFAM: PF00353; hemolysinCbind; 2.
DR PRINTS: PR00313; CABNDNGRPT.
DR PROSITE: PS00330; HEMOLYSIN_CALCITUM; 2.
SEQUENCE 998 AA; 107058 MW; 0D3BE108C309B83 CRC64;

Query Match 23.5%; Score 844; DB 2; Length 998;
Best Local Similarity 35.5%; Pred. No. 3.9e-36;
Matches 218; Conservative 91; Mismatches 185; Indels 120; Gaps 15;

QY 110 SFPPTGAKIILYIPONYDTQEGNGLOLVKAAEELGIEVOREERNINATPAOTSIGT 169
DB 25 SSSVSASAKKILLIPDNE--AAGVGINELVKADELGIHRTERRDDTAIANOFCA 81
QY 170 IOTATGLTERGIVLSAPQIDKLQK-TRAGCAL-GSASEIVONAKKATVLSIGTSLGS 227
DB 82 AEKVGVLTERGVAIPAPQIDKLQRYQVSKIGTAEVGNLAKAGIVLSALDNFTGI 141
QY 228 VLAGMDLEALQ-----NNSNOHALAKAGLELTNSLIENIANSVKTLDFEGPOISOFSGK 282
DB 142 ALSGALDELRLKQREGEDISQNDIAKSSIELINQVLDVSSINSTVFSQQLQSGF 201
QY 283 LQNIKGLGTGDKLNIGGLDRAGLDVLVSGLSGATAALVLADKNASTAKKVGAGEL 342
DB 202 LSKRRLSSVGKILGNLDPGLDGLDVSGILSAVSASFILGNSDHTGKKAAGIEL 261
QY 343 ANOVGNITKAVSYITLQRYVAAGSSTGPAVALIASTVSLAISPLARAGIADKNHAKS 402
DB 262 TTVQVGNVGNKAVSOYLQORMAQGSTTAASAGLITTSAMMLISPLFAADKFERAKQ 321
QY 403 LESYAEERKKLGVDGNLAEYORGTGTIDASVTAINALAIAGVSAAAA----- 454
DB 332 LESYERKKLTNEGDAIAGHKTGAIDAGITINIVLSSVSGVSAASASLIGAPI 381
QY 455 -----DLTFEYKHNVLITNSKKEKVTIONMF----- 483
DB 382 SMLVSLGTITSGILEASKQAMFEHVAEKFARINEMEKEHKNYFENGYDARHAFLD 441
QY 484 -----ADTAKVEPNYKARK-----DEKIEELIG--QNGERTSKOVDDLIAGKGI 530
DB 442 SLSLADVSROHAAVERAAVITHQHWDEKIGELAGITIRNADRSQ-----GK--- 487
QY 531 DELSKVYNVEYLK-HSKVNTSL-DKLISVSATSSNDNRNVLVAPTSMUDQSLSSLQ 588
DB 488 -PYIYVLENGELLEAPQEFETQVDPDQKGTIDLTSTGNVSVLTITTP----- 535
QY 589 FARGSOHWSYGLRPGSGSDWSYGLRPGSGSOHWSYGLRPGSGSDWSYGLRPGSGSOHWS 648
DB 536 -----FTPGEEVRE-----RKQSGKYEYMTSLIVNGKDIWS 566

QY 649 Y-GLRPGSGSDWS 661
DB 567 YAGIKNHKGVYDYS 580

RESULT 8
043892 PRELIMINARY: PRT: 1055 AA.
AC 043892;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE LEUKOTOXIN.
GN LKTA.
OS Actinobacillus actinomycetemcomitans
OC (Haemophilus actinomycetemcomitans).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Actinobacillus.
NCBI_TaxID=714;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-JP2;
RX MEDLINE=90202154; PubMed=2318535;
RA Kolodrubetz J., Dailey T., Kraig E.;
RT "Nucleotide sequence of the leukotoxin gene from Actinobacillus actinomycetemcomitans: homology to the alpha-hemolysin/leukotoxin gene family."
RL Infect. Immun. 58:920-929(1991).
DR EMBL: X16829; CAA34731.1; -;
DR INTERPRO: IPR001343; -;
DR PFAM: PF00353; hemolysinCbind; 3.
DR PRINTS: PR00313; CABNDNGRPT.
DR PROSITE: PS00330; HEMOLYSIN_CALCITUM; 3.
SEQUENCE 1055 AA; 113853 MW; 5331C396FA76669E CRC64;

Query Match 23.0%; Score 827; DB 2; Length 1055;
Best Local Similarity 40.1%; Pred. No. 3.3e-35;
Matches 202; Conservative 86; Mismatches 144; Indels 72; Gaps 13;

QY 114 KTGAKIILYIPONYDTQEGNGLOLVKAAEELGIEVOREERNINATPAOTSIGTOTA 173
DB 44 KTG-KKLLIYIPKN-----KKGNGLTALIKAAQKIGIEYHEGKGPPLTNGILTGTGKL 98
QY 174 IGLTERGIVLSAPQIDKLQKTR-AGQALGSASEIVONAKKATVLSIGTSLGVLG 232
DB 99 IGLTERGLTLFAPELDKWIKGKHLNSVSGTGNLTKAIDKQVSLGLQAFNLTAFSGM 158
QY 233 DLD---EALQNSN--QHALAKAGLELTNSLIENIANSVKTLDFEGEIOISOFGLONIK 287
DB 159 DLDALIKARQNKNTVDQALKASLNLINELIGTSSITNNVDFPSKQNLKIGELAGGVK 218
QY 288 GLGTGDKLNIGGLDRAGLDVLVSGLSGATAALVLADKNASTAKKVGAGFELANOV 347
DB 219 HFGSGDKLKNLPKGNLNGKIGALSVLSAISALLANNDADATATATAAAAEILTNNVL 278
QY 348 GNTKAVSYITLQRYVAAGSSTGPAVALIASTVSLAISPLAFAGIADKFNHAKSLEYA 407
DB 279 GNIGKAITQYLIQRAAAGLSTTGPVAGLIASVSLAISPLFICIAKOPBARMLEYS 338
QY 408 ERFKRLGVDGNLAEYORGTGTIDASVTAINALAIAGVSAAAA----- 454
DB 339 KRFKFGYNGSLLGQFYKNTGIDAAITTTINTVLSAIAAGVGAASAGSLVGCAPILGVS 398
QY 455 -----DLTFEYKHNVLITNSKKEKVTIONMF-----READPAKEV----- 490
DB 399 ATLSLSIGLDAKQAVYEHIANQADKIKAMENKYGANNYFENGIDARHSAFLBSLKL 458
QY 491 -----PNYKATK-----DEKIEELIG--QNGERTSKO--VDDL-----IAKNGSKI 528
DB 459 NELRREKYTENTILITQOGMDQRIQELAGITIRNGDRIGSGRAAYVDYLKKGELAKHSDKF 518


```

DR 01-AUG-1998 (TREMBLrel. 07, Created)
DR 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DR 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE ALPHA HEMOLYSIN (FRAGMENT).
GN Hyla.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ECOR 43;
RX MEDLINE=98155142; PubMed=9495754;
RA Boyd E.F., Hartl D.L.;
RT "Chromosomal regions specific to pathogenic isolates of Escherichia coli have a phylogenetically clustered distribution."
RL J. Bacteriol. 180:1159-1165(1998).
DR EMBL; AF037573; AAC38262.1; -
DR INTERPRO: IPR000911; -
DR PROSITE: PS00359; RIBOSOMAL_L11; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 181 AA; 19101 MW; A3586DC663B571E CRC64;

Query Match 10.5%; Score 377.5; DB 2; Length 181;
Best Local Similarity 45.9%; Pred. No. 6,8e-13;
Matches 83; Conservative 33; Mismatches 58; Indels 7; Gaps 3;

QY 135 GNGIADLVKAAEELIEVOEREERNIATQTSIGTITQTAIGTERGIVLSAPQIDKLQK 194
DB 1 GSSINDVLTADDELIEVOYDEKNGTAITKQVFGTAERKILGTERGVATFAFQDLKLQK 60

QY 195 -TKAGQAL-GSAESIYQANAKKATVLSGIQSLGSLVAGMDDEALQ-----NNSQNAL 247
DB 61 YQKAGNKLGSASENIGDNLKAGSVLSTFQNLGTALSSMKIDELIKKOKSGSNVSSSEL 120

QY 248 AKAGELTNSLENTANSVKTLDEFEQISQFGSKLQNKIGTIGDKLNIGGLDKAGL 307
DB 121 AKASIELINQVDTAASINNANNVNSFSQDLNKLGSVLSNTKHLNGVGNKLNLPNDNIGA 180

QY 308 G 308
DB 181 G 181

RESULT 13
ID 070070 PRELIMINARY; PRT; 181 AA.
AC 070070;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE ALPHA HEMOLYSIN (FRAGMENT).
GN Hyla.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ECOR 48, ECOL 51, ECOL 52, ECOL 54, ECOL 60, ECOL 63;
RX MEDLINE=98155142; PubMed=9495754;
RA Boyd E.F., Hartl D.L.;
RT "Chromosomal regions specific to pathogenic isolates of Escherichia coli have a phylogenetically clustered distribution."
RL J. Bacteriol. 180:1159-1165(1998).
DR EMBL; AF037579; AAC38263.1; -
DR EMBL; AF037575; AAC38264.1; -
DR EMBL; AF037576; AAC38265.1; -
DR EMBL; AF037577; AAC38266.1; -
DR EMBL; AF037578; AAC38267.1; -
DR INTERPRO: IPR000255; -

```

```

DR INTERPRO: IPR000911; -
DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
DR PROSITE: PS00359; RIBOSOMAL_L11; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 181 AA; 19104 MW; 61716DD676D6D613 CRC64;

Query Match 10.4%; Score 372.5; DB 2; Length 181;
Best Local Similarity 45.3%; Pred. No. 1.2e-12;
Matches 82; Conservative 34; Mismatches 58; Indels 7; Gaps 3;

QY 135 GNGIADLVKAAEELIEVOEREERNIATQTSIGTITQTAIGTERGIVLSAPQIDKLQK 194
DB 1 GSSINDVLTADDELIEVOYDEKNGTAITKQVFGTAERKILGTERGVATFAFQDLKLQK 60

QY 195 -TKAGQAL-GSAESIYQANAKKATVLSGIQSLGSLVAGMDDEALQ-----NNSQNAL 247
DB 61 YQKAGNKLGSASENIGDNLKAGSVLSTFQNLGTALSSMKIDELIKKOKSGSNVSSSEL 120

QY 248 AKAGELTNSLENTANSVKTLDEFEQISQFGSKLQNKIGTIGDKLNIGGLDKAGL 307
DB 121 AKASIELINQVDTAASINNANNVNSFSQDLNKLGSVLSNTKHLNGVGNKLNLPNDNIGA 180

QY 308 G 308
DB 181 G 181

RESULT 14
ID 090617 PRELIMINARY; PRT; 1729 AA.
AC 090617;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE PERICARDINE.
GN PRC.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON R;
RA Chatter A., Astler M., Zaffran S., Semeriva M., Gratecos D.;
RT "Characterization of an extracellular matrix component specifically expressed in the basal lamina of the dorsal vessel in Drosophila melanogaster."
RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF203342; AAF13280.1; -
SQ SEQUENCE 1729 AA; 164661 MW; DA9B1B7FAB5CDEDB CRC64;

Query Match 5.8%; Score 208.5; DB 5; Length 1729;
Best Local Similarity 22.2%; Pred. No. 0.014;
Matches 185; Conservative 72; Mismatches 271; Indels 307; Gaps 45;

QY 13 YGLRPGSGSQDMS---YGLRPG-----GSSQHWSTYGLRPGSGSQDMS---YGLRPG-- 57
DB 417 YGTQPGIGGTGAGQPGYGTQPGIGAGQGTGAGQPGYGTQPGIGGTGAGQPGYGTQPGIG 475

QY 58 ---GSOHWSTYGLRPGSGSQDMS---YGLRPG-----GSSQHWSTYGLRPGSGSQDMS--- 102
DB 476 YGTQPGIGGTGAGQPGYGTQPGIGAGQGTGAGQPGYGTQPGIGGTGAGQPGYGTQPGIG 534

QY 103 -YGLRPG-----GSSFP-----KTGAKKILYIPQNYQVTEGNGIADLVKAAE 146
DB 535 GTGTQPGIGAGQPGYGTQPGIGGTGAGQPGYGTQPGIGGTGAGQPGYGTQPGIGGTGAGQ 583

QY 147 ELGIEVOEREERNIATQTSIGTITQTAIGTERGIVLSAPQIDKLQKTKAGQ-ALG 202
DB 584 QPGYGTQPGVAGQGTGAGQPGYGTQPGIGGTGAGQPGYGTQPGIGGTGAGQPGYGTQPGIG 639

```

QY	203	SAESIYONANKAKT---	VLGISIOSILIGSVLAGDDLDLEALONNSNOHALAKAGIELTNSLI	259
Db	640	SOPGIGGOTGAHQPGYGTQPGVGAQGTGQPGTGAQPGI---	GGGTGAGQPG-----	688
QY	260	ENIANSYKTLDEFEEQISOFQSGSKLONIKGLGT---	LDGKL-----KNIGLIDKAG	306
Db	669	-----YRQPGIGGGQTQPGQPGYGTQGTQGTGQPGYGAQPGIGGQSGAG	736	
QY	307	L-----	GLDIYISGLISATATYALVLADKNASTAKKVGAGFELANQVYGNIT	351
Db	737	QPGYGSQPGIGGQTGQGGQPGYGSQIGGQTGA-----	GQPSYGSQPGVGAQNGGGQPGYGT	791
QY	352	KAV-----	SSYLIAQRVAAGLSSTPRVALLASTYSLAISPLRA	391
Db	792	RPVIGGGQTGAGQPGYGGQTGVGSGPGFLTPPGIGIS--	GPIGGKVGGGQSEAKPGYWA	849
QY	392	-----	GIADKFNHAKLSLESAERPKKIGDYDDNLIAEYQGTGTI---	DA 433
Db	850	QPGIGGSPRYSQPGIGDQGTGAGS-----	GYGQGPISG-QTGGGQPGYGGQA	897
QY	434	SVTAIN-----	TALAALINGYSAAADLTREKYKHNLVITNSKREKVTIONMEREAD	485
Db	898	TISGLPVGQTQPGIGALTAVPBG-----	HYGE-----TQPGIGGQTGTQNGG	940
QY	486	FAKEYPNTKATKDEKIEIILIGQNGERTTSKQVDDLI	AKNGKTTQDELISKVYDNTYELKH	545
Db	941	FGGQ-----	PGIGGQTGAG-----QPGYGFIGQPGIG-----	968
QY	546	SKNTYNSLIDKLISSVAFSTSSNDRNLVAP	TSMIDQSLSLQFANGSQHWSYGLRPGSG	605
Db	969	-----	QTGTSGRPGTGTQPGIG	986
QY	606	SQDWS---YGLRGC-----	GSSQHMWSYGLRPGSGSQ---DMSYGLRGC-----	GSQHW 647
Db	987	GQTAAGQPGYGSQPGIGGQTGACGP-GYGSQGTGCGG	IGAGQPGYGSQPGIGGQTGACGP	104
QY	648	SYGLRPGSGSQDMSYGLRPG---	GSSQHMWSYGLRPGSSQDMS---YGLRPG	693
Db	1046	GYGAGQPGYGGQGP-GYGNQPGVGQGTGAGGP-GYGSQPGYGGQGTGAGQPGYGVPIPG	1098	
RESULT 15				
Q9L469 PRELIMINARY; PRT; 1706 AA.				
AC	Q9L469			
AC	Q9L469			
DT	01-OCT-2000 (TReMBLrel. 15, Created)			
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)			
DT	01-OCT-2000 (TReMBLrel. 15, Last annotation update)			
DE	BIUNCTIONAL HEMOLYSIN-ADENYLATE CYCLASE PRECURSOR (EC 4.6.1.1).			
GN	CYAA.			
OS	Bordetella parapertussis.			
OC	Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;			
OC	Bordetella.			
OX	NCBI_taxid=519;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=63.2.			
RA	Boursaux-Bude C., Guiso N.;			
RT	"Sequencing of the Bordetella parapertussis adenylate cyclase-			
RT	hemolysin-encoding gene."			
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AJ249835; CAB76450.1; -			
KM	LysE.			
SO	SEQUENCE 1706 AA; 177038 MW; 71750E2D0BBF3B64 CRC64;			

```

Db 24 DYOHGTEQNNPPEADEKEKIVVSAATGESQMLTRGOLKEVYIGQGRGCVFEYENRABV- 352
QY 56 PGCSOHMSYGLRPGCSGSDMSYGLRPGCSSOHMSYGLRPGSGSD-----WYGLRPG 109
Db 353 -AGKSLFDPDGAAPGV-----PGGRSK-----SSPDVLETVPASGR-- 391
QY 110 SSFPTGSKKILITVPOVYDTGEOGNLOLVKAABELGIEVQREERNMIATAQTSLT 169
Db 392 ---PSLCAVE-----RQDSGIDSIDGYSRKF-----SLGE 419
QY 170 IQTALIGTERGIVLAPQIDKLQKTRAGALSAESIYONANK--AKTVLSGIQSTLGS 227
Db 420 VSDM-----AAVEAAELEMTQROVLHAGANQDDEPBGVGSASAHWGQALOGAQAIVAA 472
QY 228 VLAGMDLEALQNNNSNOHALA-----KAGIELTNSLIENIANSVKTLDFEGEDISQFS 291
Db 473 -----ORLVHAIHAIKMTQFGRAGSTNTPQBAASISAAVEFLGERSASAVALTVS 519
QY 282 KLQNKGLGTLGDKLKNTGIDRKAIGLDIVYSGILSCATFALVYADNASTAKKVGAFGE 341
Db 520 -----GFRGSSRMWAGGCVGAVGAMALGGGIAAAVAGMSLTLD--DAPAGQAAVGA 570
QY 342 LANOVVGNITVAWSYTL-----AQRVAAGLSSTGPVAAALIAVSLAISPLAFAGIADKF 397
Db 571 IALQITGTVELASISIALALAAARGVTSGLQVAGASGAANAAGALAAALSPMEITGLVQOS 630
QY 398 NHAKSLEYAERFEKLGIDGDNLLAEYQRTGCTIDASVTALINTPLAAIAGVSAAAA 454
Db 631 HVAQDLKLQAESSAYGEGALLAOLYRQDTAAEGVAVAGSALVSTVGAIVSTAAA 667

```

Search completed: March 2, 2001, 11:01:01
Job time: 471 sec

```

Query Match          5.7%; Score 205.5; DB 2; Length 1706;
Best Local Similarity 22.4%; Pred. No. 0.02;
Matches 107; Conservative 66; Mismatches 193; Indels 111; Gaps 16;

OY  %cDRSGHWSTGLRPGSGSDMSYGLRPGGSSQHWSYC-LRPGSSQ-----DMSYGLR 55
      | | | | | : : : | | | | | | | | | | | | | | | | | | | |

```


Db 80 EENNATATQTSIGTITQTAIGLTERGIVLSAQIDKLQKTRAGALSAESITVONAKAK 139
 |||||
 QY 216 TVLSGIQSILGSVLAQMDLDEALQNNNOHALAKAGLEINSLIENTIANSVKTLDEFGEO 275
 |||||
 Db 140 TVLSGIQSILGSVLAQMDLDEALQNNNOHALAKAGLEINSLIENTIANSVKTLDEFGEO 199
 |||||
 QY 276 ISQFSKLTQNTKGTGTLGDDTLKNTGGIDKAGIGDIVISGLISGATPAALVYLADKNASTAKK 335
 |||||
 Db 200 ISQFSKLTQNTKGTGTLGDDTLKNTGGIDKAGIGDIVISGLISGATPAALVYLADKNASTAKK 259
 |||||
 QY 336 VGAGELANQVYGNITTKAVSSYTLAQRVAAGLSSTGPAAALIASTVSLAISPLAFAGIAD 395
 |||||
 Db 260 VGAGELANQVYGNITTKAVSSYTLAQRVAAGLSSTGPAAALIASTVSLAISPLAFAGIAD 319
 |||||
 QY 396 KFNNAKSLESAERERKKLGVDGDNLALEYORGTGTIDASVTAINTALAIAGVSAAMA- 454
 |||||
 Db 320 KFNNAKSLESAERERKKLGVDGDNLALEYORGTGTIDASVTAINTALAIAGVSAAMAAG 379
 |||||
 QY 455 ----- 454
 Db 380 SVIASPIALLVSGITGVISTILLQYSKQAMFEHVANKIHNKIVEMKNNHGNKPYENGDA 439
 |||||
 QY 455 ----- 454
 Db 440 RLTLANLQDNMKFLNLNKLKELDAERYIAITQOQWNNICGLAGISRLGKVLSCRAYDAF 499
 |||||
 QY 455 ----- 454
 Db 500 EEGKHKADKLVQLQDSANGIIDVNSGKAKQNHILFRTPLTPTGEHNERVQTKYEYIT 559
 |||||
 QY 455 ----- 454
 Db 560 KLINRVDSWKITDGAASFTEDLTNNVORIGIELDNAGNVTKTEKYLIAKLGEQDNVF 619
 |||||
 QY 455 ----- 454
 Db 620 VGSSTTEIDGEGYDVHYSKRGNTGALTIDATKETEGSYTVNRPVETGKALHEVSTHT 679
 |||||
 QY 455 ----- 454
 Db 680 ALVGRREEKIEVRHSNNQHAGVYTKDPLKAAVEELIGTSHNDIFPKSGKFNAPFGQSDVD 739
 |||||
 QY 455 ----- 454
 Db 740 TIYGDNDGDRLEFGKGDIDLDSGNGDDEFIDGKGNLHLGKGDDIFVHRKGDGNDIITD 799
 |||||
 QY 455 -----DLTFPEKVKHNLVTTNSKKEKVTIOMNFRPADFAKEVPNTKATDE 499
 |||||
 Db 800 SDGNKLSESDSNLKDOLTFEKKHNLVTTNSKKEKVTIOMNFRPADFAKEVPNTKATDE 859
 |||||
 QY 500 KIEEIIIGNGERITRSKQVDDLIANKGNKITODELSKVVDNELLKHSKNVNSIDKLISS 559
 |||||
 Db 860 KIEEIIIGNGERITRSKQVDDLIANKGNKITODELSKVVDNELLKHSKNVNSIDKLISS 919
 |||||
 QY 560 VSAFTSSNDSRNVLVAPTSMLDQSLSLQFARGS 593
 |||||
 Db 920 VSAFTSSNDSRNVLVAPTSMLDQSLSLQFARAA 953
 |||||

RESULT 2
 A53254
 leukotoxin A - Pasteurella haemolytica (serotype T10)
 N:Alternate names: Ikta protein
 C:Species: Pasteurella haemolytica
 C:Date: 10-Aug-1990 #sequence_revision 15-Nov-1996 #text_change 18-Jun-1999
 C:Accession: S37145; A35254; S34237; S34235
 R:Linson, A.F.; Aitchison, K.; Donachie, W.
 submitted to the EMBL Data library, September 1993
 #description: DNA sequence of the leukotoxin A gene from *P. haemolytica* T10 serotype
 A:Reference number: S37145
 A:Accession: S37145
 A:Molecule type: DNA

```

A:Residues: 1955 <LAI>
A:Cross-references: EMBL:Z26247; NID:g400424; PIDN:CAAB1206.1; PID:g400425
R.Higlander, S.K.; Engler, M.U.; Weinstein, G.M.
J. Bacteriol. 172, 2343-2350, 1990
A>Title: Secretion and expression of the Pasteurella haemolytica leukotoxin.
A:Reference number: A35254; MUID:90236888
A:Accession: A35254
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 950-955 <HIG>
A:Cross-references: GB:M24197; GB:M34943; GB:M34944
R.Lainson, A.F.; Alchison, K.D.; Donachie, W.
submitted to the EMBL data library, June 1993
A:Description: DNA sequence of the carboxy terminal end of leukotoxin A from the T3 s
A:Reference number: S34235
A:Accession: S34237
A:Molecule type: DNA
A:Residues: 745-955 <LA2>
A:Cross-references: EMBL:Z22884; NID:g311828; PIDN:CAAB0498.1; PID:g311829
A:Experimental source: serotype T3
A:Accession: S34235
A:Molecule type: DNA
A:Residues: 723-955 <LA3>
A:Cross-references: EMBL:Z22887; NID:g311824; PIDN:CAAB0501.1; PID:g311825
A:Experimental source: serotype T10
C:function:
A:Description: attacks cell membranes and causes cell lysis
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: calcium binding; cytotoxicity; exotoxin; hemolysis; lipoprotein; tandem repeat
F:240-766/Domain: hemolysin A homology <HLXA>
F:718-809/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVYRF]-X)
F:718-728/Region: repeat
F:727-735/Region: repeat
F:736-744/Region: repeat
F:745-753/Region: repeat
F:754-762/Region: repeat
F:763-771/Region: repeat
F:772-780/Region: repeat
F:781-789/Region: repeat
F:792-800/Region: repeat
F:801-809/Region: repeat
F:556/Binding site: palmitate (lys) (covalent) #status predicted

Query Match          48.7%; Score 1748; DB 1; Length 955;
Best Local Similarity 43.5%; Pred. No. 3,46-77;
Matches 405; Conservative 41; Mismatches 44; Indels 442; Gaps 2;

QY      104 GLRPGSGSP-----KTKAKIILTYIPONTQYPTDEQNGGLQDLVKAEEIEGYQRER 157
           || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       24 GLNRAGQSITQAQGLTKNGAKKIILYIPKDYKYSGSGSLQDLVKAEEIEGYQKEEG 83

QY      158 NNIAATQTSISLTIOYAIGTERGIYLSAPQIDKLLOKTRKAOALGSASIVONANKAKTV 217
           || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       84 NDIAAQTSISLTIQVNLGTIERGIYLSAPQIDLKLLOKKVKQALGSSSTQNONSQAKTV 143

QY      218 LSGISIIISGLAGNDLDEALQNNNOHALAKAGIELTNSLIENIANSVTKTLDERGEQIS 277
           || | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       144 LSGVGSGNSTFYLAGNDLDEALQNESDQLTKAKGELTNLSLIENIANSVOTLDLAFSEGIS 203

QY      278 QFGSKLNIKLGTLGDKLKINIGIDKRGGLDVYISGLSGATAALVADKNASTAKRVG 337
           || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       204 QFGSKLONWKYGIALGDRLKINIGIDKRGGLGVDSRLSNTAALVLADDASTAKRVG 263

QY      338 AGFEIANDVGNITTAHSVSYLIAORVAAGLSTGPVALISTVSIALISPFAFIADKF 397
           || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       264 AGFEIANDVGNITTAHSVSYLIAQRYAAGLSTGPVALISTVAIVAISPISFAGIADKF 323

QY      398 NNAKLESIAAEFFKFLGIDGNLLAEYORGCTIDASYATANTALAAIAGGVSAANA--- 454
           || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       324 DRAKLENTAEFRFKLGEGBSLAEYOHGTCTIDASVTAINATALAAGGVSAANAGSV 383

QY      455 ----- 454
```

Db	384	VASPIALLVSGTGVISITLQYSKQAMEHVANKIHNNIVEEKKNGKNFENGYDARY	443
Oy	455	-----	454
Db	444	LANIQDNMKFLLNLNKELQAEVIAITQOOWDSNIGDLAGISRLGEKVLGKAYDAFEE	503
Oy	455	-----	454
Db	504	GOHLKADRLVOLDSAKGIIDVNTNGEAKTOHILFPTPLLPTEKREBYOTGKEYEITKL	563
Oy	455	-----	454
Db	564	HINRVDSMOIKDGAASFTFDLTNVQIRIGVELDHAENYIKTKETKVATLGDGDDNFEVG	623
Oy	455	-----	454
Db	624	SGTTEIDGEGYDRVHYSRGNYGALTIDATKETEGSGYTVNRFVBSGKALHSGTHTAL	683
Oy	455	-----	454
Db	684	VGNREEKIEYRHSHNNQHAGYTTKDTLKAVEEIIGSTHNDIPKSGKFNDAFNGDGDVDTI	743
Oy	455	-----	454
Db	744	DGNDENDRLFGGKGDIIIDGGNGDDFIDGKGNDLLHGKGGDIFVHRQDGDNDSTESE	803
Oy	455	-----DLTFEYKHNLYVTNSKKEKVTIOMMFPREDAFKEVPYKATKDEKI	501
Db	804	GNDKLSFSDSNLKDTLFEKVNHHLYITNTKQEKYTIOMMFPREAFKAKTIOMVATRDCKI	863
Oy	502	EEIIQNGERITSKQVDDLAKNGKITODELSKVVDYVELLKHKNVNTNSLDKLISVS	561
Db	864	EEIIQNGERITSKQVDELIEKNGKIAQSELTKVVDNYQLLKYRDSANSNLDKLISAS	923
Oy	562	AFTSSNDSNRNVLPATSMLODSSLQFARGS	593
Db	924	AFTSSNDSNRNVLPATSMLODSSLQFARAA	955

RESULT 3

B33389

toxin II - Actinobacillus pleuropneumoniae
N:Alternate names: cytolysin II; RTX-toxin II (ApxII)
C:Species: Actinobacillus pleuropneumoniae
C:Date: 09-Mar-1990 #sequence revision 01-Nov-1996 #text_change 18-Jun-1999
C:Accession: B33389; S18853; B43599
R:Chang, Y.F.; Young, R.; Struck, D.K.
DNA 8, 635-647, 1989
A:Title: Cloning and characterization of a hemolysin gene from Actinobacillus (Haemophilus)
A:Reference number: A33389; MUID:90126233
A:Accession: B33389
A:Molecule type: DNA
A:Residues: 1-956 <CHN>
A:Cross-references: GB:M0602; NID:g141823; PIDN:AAA67232.1; PID:g141825
A:Experimental source: serotype 5
R:Smits, M.A.; Brärlare, J.; Jansen, R.; Smith, H.E.; Kamp, E.M.; Gielkens, A.L.J.
submitted to the EMBL Data Library, July 1991
A:Description: Cytolysins of Actinobacillus pleuropneumoniae serotype 9.
A:Reference number: S18852
A:Accession: S18853
A:Molecule type: DNA
A:Residues: 1-956 <SMT>
A:Cross-references: EMBL:X61111; NID:g38939; PIDN:CAA3423.1; PID:g38941
R:Smits, M.A.; Brärlare, J.; Jansen, R.; Smith, H.E.; Kamp, E.M.; Gielkens, A.L.J.
Infected. Immun. 59, 4497-4504, 1991
A:Title: Cytolysins of Actinobacillus pleuropneumoniae serotype 9.
A:Reference number: A43599; MUID:92040145
A:Accession: B43599
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-27; 948-956 <SM2>
A:Cross-references: GB:X61111; NID:g38939

C:Comment: This organism causes porcine pleuropneumonia.
C:Genetics:
A:Gene: apxIIA; appA; clyIIIA
C:function:
A:description: attacks blood cell membranes and causes cell lysis
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: calcium binding; cytotoxic; exotoxin; hemolysis; lipoprotein; tandem repeat
F:243-787/Domains: hemolysin A homology <HLHA>
F:719-801/Region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVIRF]-X)
F:557/Binding site: palmitate (lys) (covalent) \$status predicted

Query Match 35.2%; Score 1264.5; DB 1; Length 956;
Best Local Similarity 34.3%; Pred. No. 7.8e-54;
Matches 320; Conservative 65; Mismatches 100; Indels 443; Gaps 8;

Dy 104 GLRPGSSFPKGAKKILLYIPONTQYDTDEONGLQDLVKAEEELGYEORERNINATA 163
 || | : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 34 GLTQTGHSL-QNGAKKLILTYPOG--YDSGGNGVQDLVKAAANDIGIEWREKNSNDLIA 90

Oy 164 QTSLTQITATGLPERGIVLSAPQIDKLLQKT-KAGAQLGSNESIYNQANRAKYPLSGIQ 222
 || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
Db 91 KTSFTPTQKIIGFTDRGIYLFAPQLDNLTKNPKICGNLTGSSASSISQHGKANVTLGSIQ 150

Oy 223 SLTSGVLAMGDLEALQN-NSNOHALAKAGLELNLIENIANSVKTLDEFGEIOSFGS 281
 ||||| : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
Db 151 SLTSGVLGVNINELLQKHDPQDELAKAGLETNELGNLIASSVOYDAFAEQISKLSG 210

Oy 282 KLONIKGTIGTDKKNIGLDKAGLGIYISGLISGTAALVLYADKNASTAKKVAGAFE 341
 ||||| : : : : : || : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 211 HLQNVKGJLGSNNKLONPDLGKASLGDIISGLISGSAGILLADKEASTEKKAAGE 270

Oy 342 LANQVVGNIYTVAVSSYYIAQRVAAGLSTGPAAALIASTVSIAISPLAFACIADFNFRAK 401
Db 271 FANQIIGNVTAVSSYILAQRVASGLSTGPPAALIASTVALVSPLEFLWADKFQKAD 330

Oy 402 SLESIAERFKKIDYGDNLLAEYORGCTIDASVTAINTALAIAIGVSAANA----- 454
 ::||| : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
Db 331 LKSYSERFQKLGIDGRDLADFHRETGITDASVTTINTALAIAISGVGASAGSLVAGAP 390

Oy 455 ----- 454
Db 391 VALLVAGYTGLTITLEYSKQAMPFHVANKVHDRIVEWEKKHNKNYPFOGIDSRLADLQ 450

Oy 455 ----- 454
Db 451 DNMKFLINKLELAERYVAITQQRMNDIGLAIISRDKISSKAYVDAAFEGQHQS 510

Oy 455 ----- 454
Db 511 YDSQVQDNKNGKINIISMNRKTSQVLFRTPLTPGGEENRRIOGKNSYITKLHIQRYD 570

Oy 455 ----- 454
Db 571 SWTVTDGDASSVDFNTNVQRIAVKFDAGNIEESKDTKIIANLAGANDNVFVGSSSTVI 630

Oy 455 ----- 454
Db 631 DGGGDCHDRVHSRGEYGALVIDATAETEEKSYSVKRIYGDSCALKHETTATHOTNVGNREE 690

Oy 455 ----- 454
Db 691 KLEYREDDRPHITGYTVIDSLKSVEEIIGSQPNDFFKSGQPDFDVFHGNGVDTIDGNQGD 750

Oy 455 ----- 454
Db 751 DHLFGAGDDVIDGNGNPNFLVGTGNDIISGGKDNDIYVHKTGDGNDISTDSGGODKLA 810

Oy 455 ----- 454
Db 811 FSDVMKLDITFEKKYDSSLFIINQKEKVRIGNMWFLEDLASTVANYKATNDKRKEIEIIGK 870

Oy 508 NGERTSKQVDDLAKNGKITODELSKVVDNYVELLKRSKNVTNSLDLISVSFAFTSSN 567

[illegible]

RESULT 4

toxin II - Actinobacillus suis
N:Alternate names: asha protein, cytolysin II; RTX-toxin II
C:Species: Actinobacillus suis
C:Date: 31-Dec-1993 #sequence_revision 08-Nov-1996 #text_change 05-Dec-1998
C:Accession: A43834
R:Burrows, L.L.; Lo, R.Y.
Infect. Immun. 60, 2166-2173, 1992
A:Title: Molecular characterization of an RTX toxin determinant from Actinobacillus suis
A:Reference number: A43834; MUID:92267623
A:Accession: A43834
A:Molecule type: DNA
A:Residues: 1-956 <BUR>
A:Experimental source: Isolate 3714
A:Note: sequence extracted from NCBI backbone (NCBIN:104212, NCBIPI:104211)
C:Comment: This organism causes acute fatal septicemia in young pigs.
C:Function:
A:Description: attacks cell membranes and causes cell lysis
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: calcium binding; cytotoxic; exotoxin; hemolysis; lipoprotein; tandem repeat;
F:243-787/Domain: hemolysin A homology <LUXA>
F:719-801/region: 9-residue repeats (G-G-X-G-[DN]-D-X-[LVYIF]-X)
F:557/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match	34.88;	Score 1248.5;	DB 1;	Length 956;
Best Local Similarity	34.38;	Pred. No. 4.6e-53;		
Matches 318;	Conservative 65;	Mismatches 102;	Indels 443;	Gaps 8

[illegible]

Oy	455	-----	455	-----	454
Db	571	SWTVYGDASSVDFTNVYQREAVKFPDDAGNIIESKDTKIIANLGAGNDNVFGSSTTYI	630		
Oy	455	-----	454		
Db	631	DGGGHDHVHSRGEYALYIDATAETEKSSYSVKRYVGDSKALHETIATHQNVGNREE	690		
Oy	455	-----	454		
Db	691	KIEYREDDREHGTGYTSLKSYEEIIIGSOFNDIFKGSOPDVFHGNGVDFIDNGDG	750		
Oy	455	-----	454		
Db	751	DHLEGGADVDYIDCGNGNPLVGSTGDNIIISGKDNDIYHKTGDGNDSTTDSGGODKA	810		
Oy	455	-----DLTEPKYKHLVLTINSKEKVTYIQNWFREADFAKEYPNYKATDEKIEEIIQO	507		
Db	811	FSDVNLKLFKKYDSSLEIINOGKEKVRIGMFLBEDDLASTYANKATNDRIEELIC	870		
Oy	508	NGERTSKQVDDLLAKGNGKITODELSKVYDNTVELLKHSNVNTSLDKLISVSAAFTSSN	567		
Db	871	GGERITSEOVDKLIEKNNOISALBLSKVYNDVYTSKDRONVENSJAKLISVYGSTSSS	930		
Oy	568	DSRNVL--VAFTSMLDSSLSLOPARES	593		
Db	931	DFRNNLGTYPSS--IDVS--NNIQLARAA	956		

5

toxin III - *Acinobacillus pleuropneumoniae* (serotype 8)
N:Alternate names: RTX-toxin III (AprXIII)
C:Species: *Acinobacillus pleuropneumoniae*
C:Date: 19-Dec-1993 #sequence revision 01-Nov-1996 #text change 18-Jun-1999
C:Accession: M49219; S48043; S29958
R:Jansen, R.; Briatore, J.; Kamp, E.M.; Gielkens, A.L.; Smits, M.A.
I:Infect. Immun. 61, 947-954, 1993
A:Title: Cloning and characterization of the *Acinobacillus pleuropneumoniae*-RTX-toxin
A:Reference number: M49219; MUID:93162836
A:Accession: M49219
A:Molecule type: DNA
A:Residues: 1-1052 <JANI>
A:Cross-references: EMBL:X68815; NID:G38956; PIDN:CAA8711.1; PID:G38958
A:Experimental source: strain 405, serotype 8
A:Note: sequence extracted from NCBI backbone (NCBIN:125168, NCNIP:125170)
R:Jansen, R.; Briatore, J.; Van Geel, A.B.M.; Kamp, E.M.; Gielkens, A.L.J.; Smits, M.A.
I:Infect. Immun. 62, 4411-4418, 1994
A:Title: Genetic map of the *Acinobacillus pleuropneumoniae* RTX-toxin (AprX) genes:
A:Reference number: S48042; MUID:95012630
A:Accession: S48043
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1052 <JANI>
A:Cross-references: EMBL:X80055; NID:9558150; PIDN:CAA56358.1; PID:9558152
A:Experimental source: strain 405, serotype 8
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C:Comment: This organism causes porcine pleuropneumonia.
C:Genetics:
A:Gene: aprXIIIA
C:Function:
A:Description: lyses lung macrophages
C:Superfamily: hemolysin A; hemolysin A homology
C:Keywords: calcium binding; cytotoxins; exotoxin; lipoprotein; tandem repeat; thioles
F:254-804/Domain: hemolysin A homology <HLYA>
F:736-862/Region: 9-residue repeats (G-G-X-X-G-[DN]-D-X-[LVYIF]-X)
F:571,702/Binding site: palmitate (Lys) (covalent) #status predicted

Query Match 26.6% Score 955.5 DB 1 Length 1052:
Best Local Similarity 44.1% Pred. No. 7 Ge-39:
Matches 234; Conservatively 81; Mismatches 133; Indels 83; Gaps 16;

OY	GSSPFRTAKKIIILITIPONYQVDTFQGNGIADIVAAAEELGIEVOERBRNNINATPQTSLG	168
OY	109 GSSPFRTAKKIIILITIPONYQVDTFQGNGIADIVAAAEELGIEVOERBRNNINATPQTSLG	168
Dd	48 GRAVKRYG-NKLVLVIPIK--EYDGSVGNGFPDLVAABEELGIQVKYVNNELEVAHKSLSG	104
OY	169 TLTQAIGLTERGVISAPQIDKLLOK-TRAGALGSAAE-SIVONNKKARTVLSGISIOSILG	226
Dd	105 TADQFLGLERGLTLFAHQLDQFLQKHKSISMNVGSGDAVSFKLASQSOTISSIGIOSYLIG	164
OY	227 SYLAOMDDDEALONNSNQHALAKAELLENTSILEINIANSKYTKTDFEGQISOFGSKLONT	286
Dd	165 TVLAGININEALIISGSGSELELAEGAVSYLASELVSNIAACGTTFIDAFTTOIONFGKYLENA	224
OY	287 KQLGTGLDKIKINGG-LDRKAGLGDIVISGLTGTAALVLYADKNRASTARKYGAGEELLAN	344
Dd	225 KGLGGVGRQOLNISSALSKTGLGDIITSSLISGYTASPALANKKAASSTYKAAGFEELS	284
OY	345 QVVGNITRAVSSYYLIAORVAAGLSSTGPVAALIASTVSIAISPFIAGIADKFENHKSLE	404
Dd	285 OVIGGITRAVSSYYLIAQRILAAGLSTTPGAALIASISLSAISPLAFILRVADMFNRSKEIG	344
OY	405 SYAEFRKLGVGDNDLLAEYQRCGTIDASTAIPTALAIALAGVSAAAD-----	455
Dd	345 EFAEFRKLGVGDNDLLSEFYHEACTIDASTIPTISTALSAIAGAAMAASAGALGAPIITL	404
OY	456 -----LFTEPK-----VKHNLYVTNSKKREKVLIQNWFPE-----	483
Dd	405 LVYTGITGLSIGLESKPMDHVASKIGNKIDEKKRYGKNYPENGDAYRKPALEDSE	464
OY	484 ---ADPAKEVPNYKATK----DEKIEEIIGONGERTISKQVDLLIAG-----NG	526
Dd	465 SLTSFFNKQYETERAVALLTIQQRMDEXICELAG---ITGK---GBKLSSGRAVYPDQEG	517
OY	527 KITQ---BELSVYDNVELKHKSNVTMS---LDKLSYSASFSSDSR	570
Dd	518 KLLEKKPPDFSXV--FDPTKGIEDISNSTGSTLLKFWPTL-LTPGTESR	564
RESULT 6		
S51784		
toxin III - Actinobacillus pleuropneumoniae (serotype 2)		
N:Alternate names: RTX-toxin IIIA (AprxIIIa)		
C:Species: Actinobacillus pleuropneumoniae		
C:Date: 14-Jul-1995 #sequence_revision 15-Nov-1996 #text_change 18-Jun-1999		
C:Accession: S51784		
R:Chang, Y.F.; Shi, J.; Ma, D.P.; Shin, S.J.; Lein, D.H.		
DNA Cell Biol. 12, 351-362, 1993		
A:Title: Molecular analysis of the Actinobacillus pleuropneumoniae RTX toxin-III gene cluster		
A:Reference number: S51783; MUID:93263992		
A:Accession: S51784		
A>Status: preliminary		
A:Molecule type: DNA		
A:Residues: 1-1049 <CH>		
A:Cross-references: EMBL:L12145; NID:g349605; PIDN:AAA21924.1; PID:g470685		
C:Comment: This organism causes porcine pleuropneumonia.		
C:Genetics:		
A:Gene: aprxIIIa		
C:Function:		
A:Description: lyses lung macrophages		
C:Superfamily: hemolysin A; hemolysin A homology		
C:Keywords: calcium binding; cytolysis; exotoxin; lipoprotein; tandem repeat; thiolester		
F:254-803/Domain: hemolysin A homology <HLXA>		
F:254-861/Region: 9-residue repeats (G-G-X-G-[D/N]-D-X-[L/V/I/F]-X)		
F:571,702/Binding site: palmitate (lys) (covalent) #status predicted		
Query Match 26.1%, Score 938.5; DB I; Length 1049;		
Best Local Similarity 43.5%; Pred. No. 5e-38;		
Matches 231; Conservative 80; Mismatches 137; Indels 83; Gaps 16;		
OY	109 GSFPKTKAKRIILIITIPONYQVDTFQGNGIADIVAAAEELGIEVOERBRNNINATPQTSLG	168
Dd	48 GRAVKRYG-NKLVLVIPIK--EYDGSVGNGFPDLVAABEELGIQVKYVNNELEVAHKSLSG	104

OY 169 TQIMINGTREFRIVSLAPOIDKLOK -TKAGAAIGSAP-STIYONANKKTATYSIOGISTG 226
OY 105 YADDFGLTEFLERLTLPAPLDOLFQKHSHISINWSSVSGDAHAKSLAKTSOTITISGIOSVLG 164:
OY 227 SVLAOMDDEALONNSNOHALAKAGLETNLSLENIANSVTLDEFEGSOIQFGSKLONI 286
DB 165 TYLAGININEAIISGSSELELAEAGVSILASELVSIAMACTTTIIDFTTOIONFGLAENA 224
OY 287 KCLGTUGDKLKINIG- LDRKAGLGDIVISGLSGATAALVLRADKNASTAKYVGAFELAN 344
DB 225 KQLGVGGROLOWISSASLKTGLDIITSLLSCGYRFSPALRNKA NSTKYVAAGFEELS N 284
OY 345 OVVGNSTRKAVSYIIAOPRAAGLSTGPAAIALASTSVLATISPFAACIAKPENKARLE 404
DB 285 QVIIGGITRVASYIIAORRRAGLSTTGPPAAITALIASISLASPLFLVRADFNRSKEIG 344
OY 405 SYAREFKKLGYDGDMULALEYORGCTIDASYAITMTALAAGAASAAD----- 455
DB 345 EBAERNKRKGVDKLLSFYEHAECTIDA STITTISTALSAIAGTAAMSAGALVGA PRTL 404
OY 456 -----LFTEK-----YAHNLVTNSKKEXVTIONMRE----- 483
DB 405 LVYTGITGLISGLEFSKOPMLDHVASKIGNKIDEWEKKYGKVNFENGDAYDARKKAFLEDSE F 464
OY 484 ---ADFAPKEVPNYRKPK-----DEKIEEIIQONEERTISOVDLIKK-----NG 526
DB 465 SLISTSNKYETFERVLLITOORMDEXIGBLAG----ITRG--GKLSGGARYDYFOEG 517
OY 527 KITQ---DEL SKVNDNYELLKHSKNVTNS---L DKLISSVSAFTSNSDR 570
DB 518 KLEKKPDDFSKV--FPPTKGEIDINSQNTSLTKFVTP L--LTPGRESR 564

RESULT 7
139643
RTX-toxin I - Actinobacillus pleuropneumoniae
N:Alternate names: hemolysin Apxi
C:Species: Actinobacillus pleuropneumoniae
C>Date: 19-Jul-1996 #sequence_revision 08-Nov-1996 #text_change 18-Jun-1999
C:Accession: J39643; S18769; J39645; S60732; S35781
R:Jansen, R.; Briatore, J.; Kamp, E.M.; Gielkens, A.L.; Smits, M.A.
Infect. Immun. 61, 3688-3695, 1993
A>Title: Structural analysis of the Actinobacillus pleuropneumoniae RTX-toxin I (Apix)
A:Reference number: J39641; MUID:93366425
A:Accession: J39643
A>Status: nucleic acid sequence not shown; translation not shown; translated from GB/
A:Molecule type: DNA
A:Residues: 1-1022 <RES>
A:Cross-references: EMBL:X73117; NID:g312897; PIDN:CAA51548.1; PID:g312899
R:Frey, J.; Meier, R.; Gygi, D.; Nicolet, J.
Infect. Immun. 59, 3026-3032, 1991
A>Title: Nucleotide sequence of the hemolysin I gene from Actinobacillus pleuropneuomoniae
A:Reference number: S18769; MUID:91348845
A:Accession: S18769
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-209,'AMPYLTA','218-373','R',375-561,'Q',563-686,'TC',688-1022 <PRE>
A:Cross-references: EMBL:X52895; NID:g505568; PIDN:CA48586.1; PID:g505570
R:Frey, J.; Haldemann, A.; Nicolet, J.; Bofini, A.; Prentki, P.
Gene 142, 97-102, 1994
A>Title: Sequence analysis and transcription of the apxI operon (hemolysin I) from Acqquibacterium
A:Reference number: J39644; MUID:94237497
A:Accession: J39645
A>Status: nucleic acid sequence not shown; translation not shown; translated from GB/
A:Molecule type: DNA
A:Residues: 1-209,'AMPYLTA','218-373','R',375-561,'Q',563-686,'TC',688-1022 <RES>
A:Cross-references: EMBL:X65595; NID:g505568; PIDN:CA48586.1; PID:g505570
R:Tascon, R.I.; Vazquez-Boland, J.A.; Gutierrez-Martin, C.B.; Rodriguez-Barbosa, I.; Mol, Microbiol. 14, 207-216, 1994
A>Title: The RTX haemolysins Apxi and ApxtII are major virulence factors of the swine
A:Accession: S60731; MUID:95131743

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 11:01:21 ; Search time 32.03 Seconds
(without alignments)
700.730 Million cell updates/sec

Title: US-09-306-689-13
Perfect score: 3591
Sequence: 1 MATVIDRSGHMSYGLRPGSG.....GLRPGSGSDMSYGLRPGGS 695

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2108.5	58.7	953	1	LKAL_PASHA
2	1893.5	52.7	953	1	LKAB_PASHA
3	1789	49.8	953	1	LKA3_PASHA
4	1748	48.7	955	1	LKAA_PASHA
5	1585.5	44.2	947	1	LKTA_PASSP
6	1265.5	35.2	956	1	HLXA_ACTSU
7	1264.5	35.2	956	1	RT2A_ACTPL
8	955.5	26.6	1052	1	RT32_ACTPL
9	938.5	26.1	1049	1	RT31_ACTPL
10	828.5	23.1	1023	1	RT11_ACTPL
11	828.5	23.1	1023	1	RT12_ACTPL
12	823.5	22.9	1023	1	HLX1_ECOLI
13	819.5	22.8	1024	1	HLXA_ECOLI
14	819	22.8	1050	1	LKTA_ACTAC
15	212	5.9	1705	1	CYAA_BORRE
16	206	5.7	1705	1	CYAA_BORRE
17	178	5.0	2491	1	TALA_DICDI
18	166.5	4.6	643	1	K2C1_HUMAN
19	166	4.6	638	1	K22E_HUMAN
20	166	4.6	645	1	K22E_HUMAN
21	163	4.5	622	1	KICI_HUMAN
22	157.5	4.4	507	1	ELIC_SALON
23	155.5	4.3	1045	1	GONB_CELFI
24	154	4.3	407	1	SM41_HEMPU
25	145.5	4.1	574	1	FLA3_CAMJE
26	144.5	4.0	1113	1	N116_YEAST
27	144	4.0	593	1	KICI_HUMAN
28	143.5	4.0	2869	1	RBP1_PLAYB
29	142	4.0	500	1	FLJB_SALAE
30	141	3.9	414	1	SVS2_RAT
31	140.5	3.9	933	1	SLAP_CAMJE
32	140.5	3.9	2541	1	TALI_HUMAN
33	140	3.9	682	1	PILJ_PSEAE

34	139.5	3.9	507	1	FLIC_SALBE	006968 salmonella
35	139	3.9	1509	1	MYSN_ACACA	P05659 acanthamoeb
36	138.5	3.9	995	1	Y109_YEAST	P40442 saccharomyc
37	138.5	3.9	1289	1	VG34_BPT4	P18771 bacterioph
38	138.5	3.9	1726	1	MSPI_PLAFC	P04934 plasmodium
39	138.5	3.9	1726	1	MSPI_PLAFC	P50495 plasmodium
40	138.5	3.9	2329	1	YS89_CAREL	009624 caenorhabdi
41	137.5	3.8	657	1	MCPA_CAICR	000966 caulobacter
42	137.5	3.8	2541	1	TALI_MOUSE	P26039 mus musculu
43	137	3.8	569	1	KICJ_MOUSE	P02535 mus musculu
44	136	3.8	1290	1	VACO_HELPY	P55981 helicobacte
45	136	3.8	1888	1	CAIE_CHICK	P32018 gallus gall

ALIGNMENTS

RESULT	ID	Sequence	Standard	PRT	953 AA.
1	LKAL_PASHA	01-AUG-1990 (Rel. 15, Created)			
AC	P16535	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	DE	LEUKOTOXIN FROM SEROTYPE A1.			
GN	GN	LKTA.			
OS	OS	Pasteurella haemolytica.			
OC	OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;			
CC	CC	Pasteurella.			
RN	RN	[1]			
RP	RP	SEQUENCE FROM N.A.			
RC	RC	STRAIN-SEROTYPE A1 / PHL101;			
RX	RX	MEDLINE=87306837; PubMed=3040588;			
RA	RA	Lo R.Y.C., Strathdee C.A., Shewen P.E.;			
RT	RT	"Nucleotide sequence of the leukotoxin genes of Pasteurella			
RL	RL	haemolytica A1".			
RN	RN	Infect. Immun. 55:1987-1996(1987).			
RN	RN	[2]			
RP	RP	SEQUENCE FROM N.A.			
RC	RC	STRAIN-SEROTYPE A1 / PHL101;			
RX	RX	MEDLINE=89210283; PubMed=2707120;			
RA	RA	Highlander S.K., Chidambaram M., Engler M.J., Weinstein G.M.;			
RT	RT	"Secretion and expression of the Pasteurella haemolytica leukotoxin."			
RL	RL	J. Bacteriol. 172:2343-2350(1990).			
CC	CC	"FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD			
CC	CC	CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY			
CC	CC	DERIVED.			
CC	CC	-1- SUBCELLULAR LOCATION: SECRETED.			
CC	CC	-1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING			
CC	CC	CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC			
CC	CC	ACTIVITY.			
CC	CC	-1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE			
CC	CC	INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).			
CC	CC	-1- PTR: PALMITOYLATED BY LKTC. THE TOXIN ONLY BECOMES ACTIVE WHEN			
CC	CC	MODIFIED (BY SIMILARITY).			
CC	CC	-1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.			
CC	CC	-----			
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	CC	use by non-profit institutions as long as its content is in no way			
CC	CC	modified and this statement is not removed. Usage by and for commercial			
CC	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	CC	or send an email to license@isb-sib.ch).			
CC	CC	-----			

DR	EMBL: M20730; AAA25529.1; -	
DR	EMBL: M24197; AAA25543.1; -	
DR	PIR: S29516; S29516.	
DR	HSSP: P02392; ICTF.	
DR	INTERPRO: IPR001343; -	
DR	PFAM: PF00353; hemolysincabind; 1.	
DR	PRINTS: PR00313; CABNDGRT.	
DR	PROSITE; PS00330; HEMOLYSIN_CALCULUM; 4.	
KW	Hemolysis; Toxin; Cytolysis; Cytotoxin; Repeat; Calcium;	
RV	Transmembrane; Lipoprotein; Palmitate.	
FT	TRANSMEM 229 249	POTENTIAL.
FT	TRANSMEM 297 318	POTENTIAL.
FT	TRANSMEM 366 390	POTENTIAL.
FT	DOMAIN 734 784	6 X REPEATS, GLY-RICH (BY SIMILARITY).
FT	REPEAT 734 739	1.
FT	REPEAT 743 748	2.
FT	REPEAT 752 757	3.
FT	REPEAT 761 766	4.
FT	REPEAT 770 775	5.
FT	REPEAT 779 784	6.
FT	CONFLICT 742 742	FEHVN -> LSTLQI (IN REF. 2).
FT	CONFLICT 409 414	D -> Y (IN REF. 2).
SO	SEQUENCE 953 AA; 101996 MW; 7F93D113A118C05F CRC64;	

Query Match	58.7%;	Score 2108.5;	DB 1;	Length 953;
Best Local Similarity	51.8%;	Pred. No. 1.3e-97;		
Matches 484; Conservative	1;	Mismatches 12;	Indels 437;	Gaps 2;

Oy	96	SGSOWMAYGLBPGGSSPFRTGAKKILITYPONYOYDPHOENGLDOLYKAAEELIGEYORE	155
Dd	21	SGLHAGOSLQVQASSL- KTBAKKILITIPONYOYDPHOENGLDOLYKAAEELIGEYORE	79
Oy	156	ERNNTAIQTSLSGTIQTALIGTERGIYLSAPODKLLQKTACAGALSABSIYOMANKAK	215
Dd	80	ERNNTAIQTSLSGTIQTALIGTERGIYLSAPODKLLQKTACAGALSABSIYOMANKAK	139
Oy	216	TVLSGIGSILSLVLAGMDLDEALONNSNOHLAKAGIELTNSILENTIANSEVKTIDFERGEO	275
Dd	140	TVLSGIGSILSLVLAGMDLDEALONNSNOHLAKAGIELTNSILENTIANSEVKTIDFERGEO	199
Oy	276	ISOFSKLONTKGLGTJGDKLKNIGGLDKAGLGLDVLISGLSGATAALVLADKNASTARK	335
Dd	200	ISOFSKLONTKGLGTJGDKLKNIGGLDKAGLGLDVLISGLSGATAALVLADKNASTARK	259
Oy	336	VGAGEFLANOVYGNITKAVSSYITLAORYAAGLSTSGYAAALIASTVSLAISPLAFAGIAD	385
Dd	260	VGAGEFLANOVYGNITKAVSSYITLAORYAAGLSTSGYAAALIASTVSLAISPLAFAGIAD	319
Oy	396	KFNHAKSLSESYAERPKKIGYOGDMLAEYOGSTTIDASVTAINTALAAIAGGVSAANA-	454
Dd	320	KFNHAKSLSESYAERPKKIGYOGDMLAEYOGSTTIDASVTAINTALAAIAGGVSAANA	379
Oy	455	-----	454
Dd	380	SVIASPILLLVSGITGVISTILQYSKQAMPEHVANKITHNKIVEMKNNHKNYFENGYDA	439
Oy	455	-----	454
Dd	440	RYLANI,ODNMKFLNLNKLKELOAERVIALITQOOWDNNITGLAGISRLEKVLSGRAYVDAF	499
Oy	455	-----	454
Dd	500	EEGHIKADKLVOJDSANGIITDVNSNGAKAQHLLFRTPLLTPTBEHREHYOGKYEYIT	559
Oy	455	-----	454
Dd	560	KLINRVDSMKITDGAASSTFDLTNNVQRIEILDNAGNVTKTKEFKIIAKLGGDDNVE	619
Oy	455	-----	454
Dd	620	VGSGETIEDGEGYDORVHYSKGNAGALTIDATKETTEOGSYTVNRFVETGALHEVYSTAT	679

[illegible]

RESULT	2	
LKAB_PASHA		
ID	LKAB_PASHA	STANDARD; PRT; 953 AA

DT	01-OCT-1996 (Rel. 34, Created)
DT	01-OCT-1996 (Rel. 34, last sequence update)
DT	30-MAY-2000 (Rel. 39, last annotation update)

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.

RA MEDLINE=94041617; PubMed=8225575;
Burrows L.L., Olah-Winfield E., Lo R.Y.C.;

CC -1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY

CC -1- DOMAIN: THE GLI-RICH REGION IS PROBABLY INVOLVED IN BINDING
CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
CC ACTIVITY.

CC -1- P1M: PALMITOYLATED BY LKIC. THE TOXIN ONLY BECOMES ACTIVE WHEN MODIFIED (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.

CC between the Swiss Institute of Bioinformatics and the EMBL outst
CC the European Bioinformatics Institute. There are no restrictions
CC use by non-profit institutions as long as its content is in

```
CC or send an email to license@sib.ch.
CC -----
ENR1 10101E 000000 1
```

DR PRINTS; PR00313; CABNDNGRPT.
DR PROSITE; PS00330; HEMOLYSIN_CALCITUM; 4.

FT	TRANSMEM	297	317	POTENTIAL.
FT	TRANSMEM	381	401	POTENTIAL.

FT	TRANSMEM	POTENTIAL.
FT	381	6 X REPEATS, GLY-RICH.
DOMAIN	784	

QY	158	NNIAATGSLGTIGTATIGTETRGIVLSPAPDILKJTKKQAGLGAESLIVONANKKTV	217
Db	82	NDIAKQSLSTIGVNLGTLTRGIVLSPAPDILKJTKKQAGLGAESLIVONANKKTV	1411
QY	218	LSGIOSLIGSVLAGMDDEALQONNSNOHALAKAGLELTNSLIENIANSVTLDEFFGHOIS	2777
Db	142	LSGVOSLIGSVLAGMDDEALQONNSDQTLTAKAGLELTNSLIENIANSVOTLDAFSPQIS	201
QY	278	QFGSKLQWIKGLTGLGKLNKIGGLDKRAGLGLVYISGLSGATATVLAQKMASTARKVG	337
Db	202	QFGSKLQWIKGLGALGDLKLNKIGGLDKRAGLGLVHISGLSGATATVLAQKMASTARKVG	261
QY	338	AGFELANOVNITRKAVSSYTLAQRVAAGLSSTGPVALASTVSLASPLAFAGINDKF	3979
Db	262	AGFELANOVNITRKAVSSYTLAQRVAARLSSTGPVALASTVSLASPLAFAGINDKF	321
QY	398	NHANSLESYAEERFKKLGVDGNLLAEYQRTGTTIDASVTALNTALAAIAGVSAASAAA	4544
Db	322	DRAKSLERYAEERFKKLGEGDLSLAEYQHGCTGTIDASVTALNTALAAIAGVSAASAAASV	381
QY	455	-----	454
Db	382	VASPIALLVSGITGVISTITIQYSKOAMFEHYANKINHKIVEMEKNGGKNYPENGYDARY	4411
QY	455	-----	454
Db	442	LANIQDNMKFLNLNKLQAEARVLAITQQQWDSNIGPLAGISRLGEVLSGKAYVDAFEE	501
QY	455	-----	454
Db	502	GOHLKADKLVLQDSAKGIIDVSNTEGAKTQHILFRPLTPGTETKRREROVTKREYETTKL	561
QY	455	-----	454
Db	562	HINRYDSWQIKDGAASSTFDLTNNVQRIQVELDHAENVIKTKETKIYATLGDGDNNFVG	621
QY	455	-----	454
Db	622	SGTTEIDGEGEYDVAHYSRQNYGALLTIDATKEPRQSGYTVNRFESGKALHEVTSHTAL	681
QY	455	-----	454
Db	682	VGNREKIEYHRSHNNQHHAGYTRTKDLKAVEEIIIGTSHNDIFGSKFNDAPNGDGYDTI	741
QY	455	-----	454
Db	742	DQNDGNRLFGGKGDIIIDGNGEDPFDGKGKNNLLHGKGGVDFYHROGDQNDSTIESE	801
QY	455	-----DLTEKYKHNLVITNSKREKVTIONMFEADFAKEVBNYKATDEKI	501
Db	802	GNDKLSFDSNMLKDLTEKYKHNLVITNTKQEKVTIONMFEADFAKTIONVYATRODKI	861
QY	502	EEIIGQNGERTTSKQVVDLLAKGNGKITODELSKVYDNYNELKHSKAVTNSLDKLISVS	561
Db	862	EEIIGQNGERTTSKQVVDLLAKGNGKIAQSELTQVYDNYDOLKYSRPAASNSLDKLISVS	921
QY	562	AFTSSNDSRNVLAFTSMQDSLSLQFARQS	593
Db	922	AFTSSNDSRNVLAFTSMQDSLSLQFARQA	953
RESULT 4			
ID	LKAA_PASHA	STANDARD:	PRT: 955 AA.
AC	P5117:		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DT	30-MAY-2000 (Rel. 39, Last annotation update)		
DE	LEUKOTOXIN FROM SEROTYPE T10.		
OS	GN		
OC	Pasteurella haemolytica.		
	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;		

```

CC Pasteurella.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SEROTYPED T10.
RX MEDLINE=96425875; PubMed=8828217;
RA Latson F.A., Murray J., Davies R.C., Donachie W.;
RT "Characterization of epitopes involved in the neutralization of
RL Pasteurella haemolytica serotype A1 leukotoxin."
RT Microbiology 142:2499-2507(1996).
CC -1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC DEFINED.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
CC ACTIVITY.
CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE
CC INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).
CC -1- PTM: PALMITOYLATED BY LKTC. THE TOXIN ONLY BECOMES ACTIVE WHEN
CC MODIFIED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z26347; CAAB1206.1; -
DR INTERPRO: IPR001343; -
DR PFAM: PF00353; hemolysinCdbind; 1.
DR PRINTS: PR00313; CABNDNGRPT.
DR PROSITE: PS00330; HEMOLYSIN_CALCIUM; 4.
KW Hemolysis; Toxin; Cytolysis; Cytotoxin; Repeat; Calcium;
KW Transmembrane; Lipoprotein; Palmitate.
FT TRANSMEM 299 319 POTENTIAL.
FT TRANSMEM 381 381 POTENTIAL.
FT TRANSMEM 361 403 POTENTIAL.
FT DOMAIN 736 786 6 X REPEATS, GLY-RICH.
FT REPEAT 736 741 1.
FT REPEAT 745 750 2.
FT REPEAT 754 759 3.
FT REPEAT 763 768 4.
FT REPEAT 772 777 5.
FT REPEAT 781 786 6.
SQ SEQUENCE 955 AA; 102187 MW; B60F2DB8168BCAF CRC64;

Query Match 48.7%; Score 1748; DB 1; Length 955;
Best Local Similarity 43.5%; Pred. No. 1e-79;
Matches 405; Conservative 41; Mismatches 44; Indels 442; Gaps 2;

QY 104 GLRPGSSFP-----KTKAKKIILYIPONTQYPTDEONGLQDLYVKAEEIGYEQREER 157
Db 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
24 GLNRAGQSLTPAQGTILNKKGAKKIILYIPKDKYKSGSGNGLQDLYVKAEEIGYEQREEG 83

QY 158 NNMTATQTSLSGTOTATGILTERGIVLSPAPQDKLQTKAGQALGSASIVQANKAKTV 217
Db 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
84 NDIKAQTSLSLTIONVGLTERGIVLSPAPQDKLQTKAGQALGSASIVQANKAKTV 143

QY 218 LSGIOSTLGSVLAMDMDDEALQNNNSNOHALAKAGIELTNSLIENINANSVKTLDEGEQIS 277
Db 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
144 LSGVQGSNRYVLAMGMDLDEALQNNSDQLTLAKAGIELTNSLIENINANSVQTLDAFSDQIS 203

QY 278 QFGSKLQNIKIGTLGDKLKNIGGIDKAGLGLDVIYISGLSGATFALVADKNASTAKKRV 337
Db 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
204 QFGSKLQNVKCLGALGDKLKNIGGIDKAGLGLDVKSRLLSGATFALVADKDASTAKKRV 263

QY 338 AGFELANQVGNITKRAVSSYLLAQRVAAGLSSTGCPVAALLASTVSLAISPAPGIADKF 397
Db 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
264 AGFELANQVGNITKRAVSSYLLAQRVAAGLSSTGCPVAALLASTVSLAISPAPGIADKF 323

```

```

OY 398 NNAKSLAESAEERFKKLGVDGNLAEYORGSTIDASVTAINTALAIAGVSAANA----- 454
DB 324 DAKSLAENAEERFKKLGVDGNLAEYORGSTIDASVTAINTALAIAGVSAANAAGSV 383
OY 455 ----- 454
DB 384 VASPIALLVSGITGYISTILQYSKQAMFEHVANKITHNKIVEMKNNCKNFENGVDARY 443
OY 455 ----- 454
DB 444 IANLQDNMFKLLNKLQAEERVAITQOQMDNIGDLAISRLEGKVLGKRAYDAFEE 503
OY 455 ----- 454
DB 504 GOHLKADKLVDLSAKGIIDVTNTGSAKTQHLFPRLTPTGTERKERVQTSKEYITKL 563
OY 455 ----- 454
DB 564 HINRVDSMOKKGAASFTFDLINVQRIQVELDHAENVIKETKIVATLGDGDDNVFVG 623
OY 455 ----- 454
DB 624 SGTTEIDGEGYDRVHYSRGNYGALTIDATKEEGSYTVNRFVESGKALHGTSTHTAL 683
OY 455 ----- 454
DB 684 VGNREEKLEIYRSHNNQHNAGYVTKDTLKAVEEIICTSHNDIFKSGKFNDAFNGGQVDTI 743
OY 455 ----- 454
DB 744 DQDNDNRLFGCKGDDIIDGNGDDEFIDGCKNDLHGKGDIDIVHROGDGNDSTTSE 803
OY 455 -----DLFEKVKHNLVITNSKKEKVTIQQNMFREDAFEKVPNKATKDEKI 501
DB 804 GNDKLSFSDSNLKDLEFEKVNHLVITNTKQEKVTIQNMFREDAFEKTIQNVATRDCKI 863
OY 502 EEIIONCEERINSKOVDDLIAGKNGKITODELSKYVDNVEILKHSKNVTNSIDKLSSS 561
DB 864 EEIIONCEERINSKOVDDLIAGKNGKITODELSKYVDNVEILKHSKNVTNSIDKLSSS 561
OY 562 AFTSSNDRSNVLVAPTSMLDQSLSSLOFARGS 593
DB 924 AFTSSNDRSNVLVAPTSMLDQSLSSLOFARGS 955

RESULT 5
ID LKTA_PASSP STANDARD: PRF: 947 AA.
AC P55123.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE LEUKOTOXIN.
GN LKTA.
OS Pasteurella haemolytica-like sp. (strain 5943B).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93239320; PubMed=8478098;
RA Chang Y.-F., Ma D.-P., Shi J., Chengappa M.M.;
RT "Molecular characterization of a leukotoxin gene from a Pasteurella
RT haemolytica-like organism, encoding a new member of the RTX toxin
RT family."
RL Infect. Immun. 61:2089-2095(1993).
CC -I- FUNCTION: VIRULENCE FACTOR WHICH IS CYTOTOXIC FOR LEUKOCYTES BUT
CC IS NOT HEMOLYTIC.
CC -I- SUBCELLULAR LOCATION: SECRETED.
CC -I- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
CC ACTIVITY.
CC -I- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE

```

```

CC INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).
CC -I- PFM: PALMITOYLATED BY LKTC. THE TOXIN ONLY BECOMES ACTIVE WHEN
CC MODIFIED (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
DR EMBL: L12148; AAA16444.1; -.
DR INTERPRO: IPR001343; -.
DR PFM: PF00353; HemolysinCabin; 1.
DR PRINTS: PR00313; CABENDGRPT.
DR PROSITE: PS00330; HEMOLYSIN_CALCIUM; 4.
KM Hemolysis; Toxin; Cytolysis; Cytotoxin; Repeat; Calcium;
KM Transmembrane; Lipoprotein; Palmitate.
FT TRANSMEM 154 170
FT TRANSMEM 312 333
FT TRANSMEM 393 414
FT DOMAIN 625 780
FT REPEAT 625 630
FT REPEAT 730 735
FT REPEAT 739 744
FT REPEAT 748 753
FT REPEAT 757 762
FT REPEAT 766 771
FT REPEAT 775 780
SQ SEQUENCE 947 AA; 101559 MW; 9744F06395EF5BED CRC64;

```

Query Match 44.2%; Score 1585.5; DB 1; Length 947;
 Best Local Similarity 39.7%; Pred. No. 1.1e-71;
 Matches 370; Conservative 56; Mismatches 70; Indels 437; Gaps 3;

```

OY 96 SGGQWSYGLRPGSSFPKTKAKIILYIPONYQYDEOGNGLDVLKAEELGIEVORE 155
DB 17 SGLHKTGOSLQAGOSL-KAGAKKILLYIPKDYEDSGRGNLODLVRAAEDLGIEVORE 75
OY 156 EENNATATOTSIGTQTAIGLTERGIVASAPQIDKLOKTRAGQALGSAESIYQANAK 215
DB 76 ERNGIATQNSLSTIQNLIGSESEGVVLSAPQDLKLOKTRAGQALGSAESIYQANAK 135
OY 216 TVLSGIQSLGVLGNDLDEALQNSNOHALAKAGLELNSLJENINSVKTLDPEEQO 275
DB 136 TLLSGIQLSGVMAAGMDLDELKKGSELDLAKAGLELNSLJENINSVKTLDPEEQO 195
OY 276 ISQFGSKLQNIKGLTGLDKLKNIGLIDKAGLGLDIVISGLLSGATPAALVLADKNASTAK 335
DB 196 ISQFGSKLQNIKGLTGLDKLKNIGLIDKAGLGLDIVISGLLSGATPAALVLADKNASTAK 255
OY 336 VAGFELANQVGNITKAVSSYIILQRYAAGLSSGPVAAALIASVLSAISPFAAGIAD 395
DB 256 VAGFELANQVGNITKAVSSYIILQRYAAGLSSGPVAAALIASVLSAISPFAAGIAD 315
OY 396 KFNNAKSLAESAEERFKKLGVDGNLAEYORGSTIDASVTAINTALAIAGVSAANA- 454
DB 316 KFNNAKSLAESAEERFKKLGVDGNLAEYORGSTIDASVTAINTALAIAGVSAANAAG 375
OY 455 ----- 454
DB 376 SLVGAPIALLVSGITGYISTILQYSKQAMFEHVANKIHDKIYDMEKHNKKNFENGYS 435
OY 455 ----- 454
DB 436 RYLADLQDNMROKLNKLQAEERVAITQOQMDNIGDLAISRLEGKVLGKRAYADAF 495
OY 455 ----- 454
DB 496 EEGKIKADTFVQDLSATGVINTSKSDNVKTOHILFPTPLTPGVENNERIQTGKEYITT 555

```

QY 455 ----- 454
 Db 556 KLINIRVDSMKITDGAITNSFDLNNVQRIEILDHADNVTIKETKILIANLGDNDF 615
 QY 455 ----- 454
 Db 616 IGSQTEVDGNGIDRVHYSRGDYGALITDATNESVGSYTKRFVETGALHEVTATQS 675
 QY 455 ----- 454
 Db 676 VLVGSREKIEIHRSSNTOHAGIYTTDTLKSVEEIICTSRNDLFKSGKFDPAFHGDGVD 735
 QY 455 ----- 454
 Db 736 NIDGNAGNDRLFSGKGFIDIDGGGDGDFIDGGGDDILHGKGNLICTYKGNDSIDS 795
 QY 455 ----- 500
 Db 796 GGNDRLSFADSNLKDLEFEKVNHLMTNVTYKKEKVTIONFREADYAKTYHNTQATADEK 855
 QY 501 IEEITGNGERTSKOYDDIAGKNGKITODELSKYVDNELKHSNVTNSIDKLISV 560
 Db 856 IEEITGNGERTSKQIDELIEKKGKITDSELEKRIASSALLESKFASNSLKLTVSA 915
 QY 561 SAFTSSNDSRNVLVAPTSMLODLSLQFARGS 593
 Db 916 GAFASSNDNRVGLGVPSTLYEHT-QSVQFVRAA 947

RESULT 6

HLVA_ACTSU STANDARD; PRT; 956 AA.
 ID 000951;
 AC 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE HEMOLYSIN (CYTOLYSIN II) (CLY-IIA) (HLX-IIA) (CYTC) (APPA).
 GN APPA OR CLYIIA OR HLXIIA OR CYTC.
 OS Actinobacillus suis.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Actinobacillus.
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-3714;
 RX MEDLINE-92267623; PubMed-1587585;
 RA Burrows L.V., Lo R.Y.;
 RT "Molecular characterization of an RTX toxin determinant from
 Actinobacillus suis."
 RL Infect. Immun. 60:2166-2173(1992).
 CC -1- FUNCTION: ONE OF THE VIRULENCE FACTORS OF A.SUIS. MIGHT BE A
 SECRETED CYTOTOXIN, POSSIBLY THE EXTRACELLULAR HEMOLYSIN.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
 CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
 ACTIVITY.
 CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE
 INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).
 CC -1- PTM: PALMITOYLATED BY LKTC. THE TOXIN ONLY BECOMES ACTIVE WHEN
 MODIFIED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation at
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: M90440; AAA21918.1;
 DR INTERPRO: IPR001343;
 DR PFAM: PF00353; hemolysinCbind; 1.

DR PROSITE: P500330; HEMOLYSIN_CALCIUM; 1.
 KW Hemolysis; Toxin; Cytolysis; Cytotoxin; Repeat; Calcium;
 KM Transmembrane; Lipoprotein; Palmitate.
 FT TRANSMEM 238 254 POTENTIAL.
 FT TRANSMEM 302 320 POTENTIAL.
 FT TRANSMEM 383 406 POTENTIAL.
 FT DOMAIN 719 799 9 x REPEATS, GLY-RICH.
 FT REPEAT 719 724 1.
 FT REPEAT 728 733 2.
 FT REPEAT 737 742 3.
 FT REPEAT 746 751 4.
 FT REPEAT 755 760 5.
 FT REPEAT 764 769 6.
 FT REPEAT 773 778 7.
 FT REPEAT 782 787 8.
 FT REPEAT 794 799 9.
 SO SEQUENCE 956 AA; 102453 MW; 3415FF1D7AD4365 CRC64;

Query Match 35.2%; Score 1265.5; DB 1; Length 956;
 Best Local Similarity 34.5%; Pred. No. 8.4e-56;

Matches 320; Conservative 65; Mismatches 100; Indels 443; Gaps 8;

QY 104 GLRPGSSFPKGTAKKILYIPONYOYDTBOGNGLODLYKAAELGIEVOREERNIATA 163
 Db 34 GLNOTGSHL-QNGAKKILYIPOG--YDSGQNGIIDLKAAADLGIEVREBRNSLIDIA 90
 QY 164 QTSIGTQTAIGTERIVTASPOIDKLOKT-KAGALGSARSYONAKATVLSGTO 222
 Db 91 KTSFDTYQKILGFTDRKIVTAPOLDNLKPKRIGTGLTSSASSISQNGKAMTVLGGIO 150
 QY 223 SILGSVLGMDLDEALON-NSNOHALAKAGLELNTSLIENIASVYTLDFEGQISOFGS 281
 Db 151 SILGSVLGVNVLNELLONKPNQDELAKAGLELNTSLIENIASVQTVDAFAQISKLS 210
 QY 282 KIONKGLGTGLKLNKIGGLKAGLGLDVISGLSGATAVLADKNASTARKVAGFE 341
 Db 211 HLONVKGLGLSKLNQMLPOLGKASIGLIDTISGLSGASGLILADKKASTEKKAAGVE 270
 QY 342 LANOVGNITKRAVSVYTLAORVAAAGSSTGPVAAALIASVSLAISPLAFAGIDKKNHAK 401
 Db 271 PANQITGNVTKRAVSVYTLAORVAAAGSSTGPVAAALIASVSLAISPLFNVADKFRQAD 330
 QY 402 SLESYAEERFKLGYDGDNLAEYORGTTIDASVTATNTALAIAGVSAANA----- 454
 Db 331 LKISYSERFQKLGTDGRLLADPHRETGITDASVTITNTALAIAGVGAASGLVGAP 390
 QY 455 ----- 454
 Db 391 VALLVAGVTGLITILEYSKQAMFEHYANKHVDRIWEMERKHNKYPEQGYDSRHLADQ 450
 QY 455 ----- 454
 Db 451 DNMFELINKLEOAEERVAITQQRMDNIGDLAISRTQTSKGAAYDAFEGNISP 510
 QY 455 ----- 454
 Db 511 SHIPYSIDNKNGLIINISNTKRTQSVLFTPLLTPEGENERTQEQKNSYITLHLQNDV 570
 QY 455 ----- 454
 Db 571 SMTVTVGADASSVDFTNVVQRIAVKFFDAGNIIESKDTKILIANLGAGNDNVFSGSSTVYI 630
 QY 455 ----- 454
 Db 631 DGGDGDHVRHYSRGEYALVIDAFETEGSGYSVKRYVGDSCALHETIATHQTVGNREE 690
 QY 455 ----- 454
 Db 691 KIEYREDDRFHTGTYVTSLSKSVIELISQFNDIFKSGQFDVPHGANGVDTIDGNDGD 750
 QY 455 ----- 454

Db 751 DFLFGAGDDVDIDGNGNFWLGTGNDIISGKNDIYVHKTGNDISITDGGQDKLA 810
QY 455 -----DLTEPKVKNHNVITNSKKEKTYTQWMPREADFAKVPYKATKDKIEIIGQ 507
Db 811 FSDVWLKDLTEPKVSSLEIINQKGEKVRIGMFLLEDIASTVAVYKAKNDRIEIIIGK 870
QY 508 NGERITSKQVDDLIKAKNGKITODELSKVVDNYELKSHKNYTNLSDKLISVSASFSSN 567
Db 871 GGERITSEQVDLIKEGNQIISAEALSKVYNDYNTSKDRQNVSNLSLAKLISSVGSFTSSS 930
QY 568 DSRNVL--VAPTSLMDQSLSSIQFARGS 593
Db 931 DFRNMIGTYVPSS-IDVS-NNIQLARAA 956
RESULT 7
RT2A.ACPL STANDARD: PRT: 956 AA.
AC P15377:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RTX-II TOXIN DETERMINANT A (APX-IIA) (HEMOLYSIN IIA) (HLY-IIA)
DE (CYTOLYSIN IIA) (CLY-IIA).
GN APXIIA OR CLYIIA OR HLYIIA OR APPA OR CYTC.
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae.
OC Actinobacillus.
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-SEROTYPE 5;
RX MEDLINE=90126233; PubMed=2693022;
RA Chang Y.-F., Young R., Struck D.K.;
RT "Cloning and characterization of a hemolysin gene from Actinobacillus
RL DNA 8:635-647(1989).
[2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN-ISOLATE CVI 13261 / SEROTYPE 9;
RX MEDLINE=92040145; PubMed=1937809;
RA Smits M.A., Briatore J., Jansen R., Smith H.E., Kamp E.M.,
RA Gielkens A.L.;
RT "Cytolysins of Actinobacillus pleuropneumoniae serotype 9.";
RL Infect. Immun. 59:4497-4504(1991).
CC -I- FUNCTION: ONE OF THE VIRULENCE FACTORS OF A.PLEUROPNEUMONIAE,
CC WHICH SHOWS A WEAK HEMOLYTIC ACTIVITY AND IS MODERATELY CYTOTOXIC
CC FOR ALVEOLAR MACROPHAGES AND NEUTROPHILS.
CC -I- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
CC ACTIVITY (BY SIMILARITY).
CC -I- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE
CC INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).
CC -I- PTM: PALMITOYLATED BY APXIIIC. THE TOXIN ONLY BECOMES ACTIVE WHEN
CC MODIFIED (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M30602; AAA87232.1; -;
DR EMBL: X61111; CAA43423.1; -;
DR PIR: B33389; B33389;
DR PIR: S18853; S18853;
DR INTERPRO: IPR001343; -;
DR PFAM: PF00353; hemolysinCabinid; 1.
DR PROSITE: PS00330; HEMOLYSIN_CALCIUM; 1.
KW Hemolysin; Toxin; Cytolysin; Repeat; Calcium;

KW Transmembrane; Lipoprotein; Palmitate.
FT TRANSMEM 233 256
FT TRANSMEM 266 323 POTENTIAL.
FT TRANSMEM 361 406 POTENTIAL.
FT DOMAIN 719 787 8 X REPEATS, GLY- RICH.
FT REPEAT 719 724 1.
FT REPEAT 728 733 2.
FT REPEAT 737 742 3.
FT REPEAT 746 751 4.
FT REPEAT 755 760 5.
FT REPEAT 764 769 6.
FT REPEAT 773 778 7.
FT REPEAT 782 787 8.
SQ SEQUENCE 956 AA; 102531 MW; BDBCABBAADF14A641 CRC64;
Query Match 35.2%; Score 1264.5; DB 1; Length 956;
Best Local Similarity 34.5%; Pred. No. 9,4e-56;
Matches 320; Conservative 65; Mismatches 100; Indels 443; Gaps 8;
QY 104 GURPGSSFPPTGAKKILLYIPONYOTDEONGIQLDLYKAAEELIEVORENNIATYA 163
Db 34 GLTQGHSL-ONGAKKILLYIPQG--YDSGGNGVQDLVKAAANDGIEVMEERSNDLIA 90
QY 164 QTSIGTQATGTPRGIVLSAPQIDKLQRT-KAGQALGSSESTYONAKAKTYLSGIQ 222
Db 91 KTSFDTYKILGFDRGIVLPAPODLNLKKNPKIGNTLGSASSISQNIIGKANTVLSGIQ 150
QY 223 SLTGSVLGMDLDEALQN-NSNOHALAKAGLETNLSLENIANSVKTDEFGEOISPGS 281
Db 151 SLTGSVLGSVGNLNLQKKDPQOLELAKAGLETNLSLVGNLIASSQYDAFAEOISKLSG 210
QY 282 KLQNIKGLGTGDKIKNTIGDKAGLGDVLSGLISGATVLAALVADKNASTAKKYGAGE 341
Db 211 HLQNKVGGLGSNNKLNQNPDKKASLIGDIIISGLSGASAGILADKEASTKKAAAGE 270
QY 342 LANOVGNITAVASSYILAOVPAAGLSTGPAALIASVLSLSPAFAGIADKFNHAK 401
Db 271 FANQITIGVTVAVSSYILAOVPAAGLSTGPAALIASVLSLSPAFAGIADKFNHAK 330
QY 402 SLESAERFKKIGYDGNMLAEYQRTGTIDASVATNTALAAIAGVSAAA----- 454
Db 331 LKKSSERFQKIGYDGNMLAEYQRTGTIDASVATNTALAAIAGVSAAA----- 390
QY 455 ----- 454
Db 391 VALLAVGTGLITLLEYSKQAMFEHVANKYHRIWEMKKHNKXFPQGYDSRHLADLQ 450
QY 455 ----- 454
Db 451 DNMKFLINLKEQLAERVVAITQQRMDNIGDLAISRRTDKISSGKAYVDAFEGOHQS 510
QY 455 ----- 454
Db 511 YDSSVQDNKNGIINISMTNKTQSVLFRTPLTPGGEENRERIQGRKNSYITKLHIQVD 570
QY 455 ----- 454
Db 571 SWTVTDGADASSVDFTNVVORLAVKFDAGNIIESKQRIIANLACAGDNVFGSSTVI 630
QY 455 ----- 454
Db 631 DGGDGHDRVHYSRGEYALVIDATAETEKGSYKRVYVGSKALHETTATHTQNVGNREE 690
QY 455 ----- 454
Db 691 KIEYRREDRHTGTGTYTDSLSKVEELIGSQFNDIFKRSQFDDVPHGNGVDTIDGNGCD 750
QY 455 ----- 454
Db 751 DFLFGAGDDVDIDGNGNFWLGTGNDIISGKNDIYVHKTGNDISITDGGQDKLA 810
QY 455 -----DLTEPKVKNHNVITNSKKEKTYTQWMPREADFAKVPYKATKDKIEIIGQ 507

```

Db      811 FSDVNLKDLTFKRVSDSLFEIINQKGEKVRIGNMFLEDDLSTVANYKATDRKIEETIGK 870
OY      508 NGRIRTSKQVDDLIANKNGKITODELSKVVDNELLKHSNVNVLKLSVSAFTSSN 567
Db      871 GGEIRTSQVDDKLTKEGNNQISAEALSKVVDNVTSTKRONVSNLAKLSSVGSFTSSS 930
OY      568 DSRNVL--VAPTSMDQSLSSLOFARGS 593
Db      931 DFRNNGTVPSS-IDVS-NNIQLARAA 956

RESULT 8
RT32_ACTPL STANDARD; PRT; 1052 AA.
ID      RT32_ACTPL STANDARD; PRT; 1052 AA.
AC      P55131;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DE      30-MAY-2000 (Rel. 39, Last annotation update)
DE      RTX-III TOXIN DETERMINANT A FROM SEROTYPE 8 (APX-IIIA) (CYTOLYSIN
DE      IIA) (CLY-IIIA).
GN      APXIIA OR CLYIIIA OR RTX A OR PTXA.
OS      Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC      Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC      Actinobacillus.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-405 / SEROTYPE 8;
RX      MEDLINE-95012630; PubMed-7927703;
RA      Jansen R., Briatore J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J.,
RA      Smits M.A.;
RT      "Genetic map of the Actinobacillus pleuropneumoniae RTX-toxin (ApX)
RT      operons: characterization of the ApXIII operons.";
RL      Infect. Immun. 62:4411-4418(1994).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN-SEROTYPE 8;
RX      MEDLINE-93162836; PubMed-8432615;
RA      Jansen R., Briatore J., Kamp E.M., Gielkens A.L.J., Smits M.A.;
RT      "Cloning and characterization of the Actinobacillus
RT      pleuropneumoniae-RTX-toxin III (ApXIII) gene.";
RL      Infect. Immun. 61:947-954(1993).
RN      [3]
RP      FUNCTION: DOES NOT HAVE HEMOLYTIC ACTIVITY BUT SHOWS A STRONG
RN      CYTOTOXICITY TOWARDS ALVEOLAR MACROPHAGES AND NEUTROPHILS.
RN      [4]
RP      SUBCELLULAR LOCATION: SECRETED.
RN      [5]
RP      DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
RN      CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
RN      ACTIVITY (BY SIMILARITY).
RN      [6]
RP      -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE
RN      INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).
RN      [7]
RP      -1- PTM: PALMITOYLATED BY APXIII. THE TOXIN ONLY BECOMES ACTIVE WHEN
RN      MODIFIED (BY SIMILARITY).
RN      [8]
RP      -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
RN      [9]
RP      This SWISS-PROT entry is copyright. It is produced through a collaboration
RN      between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN      the European Bioinformatics Institute. There are no restrictions on its
RN      use by non-profit institutions as long as its content is in no way
RN      modified and this statement is not removed. Usage by and for commercial
RN      entities requires a license agreement (See http://www.isb-sib.ch/announce/
RN      or send an email to license@isb-sib.ch).
RN      [10]
RP      EMBL: X80055; CAA56358.1; -
RN      EMBL: X68815; CAA48711.1; -
RN      INTERPRO: IPR001343; -
DR      PFM: PFM00353; hemolysinCabdnd; 2.
DR      PROSITE: PS00330; HEMOLYSIN-CALCIUM; 3.
KW      Toxin; Cytolysis; Hemolysin; Repeat; Calcium; Transmembrane;
KW      Lipoprotein; Palmitate.
FT      *TRANSMEM 248 265
FT      *TRANSMEM 275 334
FT      *TRANSMEM 372 418
FT      *DOMAIN 754 859
FT      POTENTIAL.
FT      POTENTIAL.
FT      7 X REPEATS, GLY-RICH.

```

```

FT      REPEAT 754 759 1.
FT      REPEAT 763 768 2.
FT      REPEAT 772 777 3.
FT      REPEAT 781 786 4.
FT      REPEAT 790 795 5.
FT      REPEAT 799 804 6.
FT      REPEAT 808 813 7.
FT      REPEAT 827 832 8.
FT      REPEAT 836 841 9.
FT      REPEAT 845 850 10.
FT      REPEAT 854 859 11.
SQ      SEQUENCE 1052 AA; 112809 MW; F83AFE25A6FD8758 CRC64;

Query Match 26.6%; Score 955.5; DB 1; Length 1052;
Best Local Similarity 44.1%; Pred. No. 2.2e-40;
Matches 234; Conservative 61; Mismatches 133; Indels 83; Gaps 16;

OY      109 GSSEPKTGAKRIILYIPONYQYDTEDGNGLODLYKAAELGIEVOREERNNTAOTSIG 168
Db      48 GKAVQKYG-NKVLVLRK--EYDGSVGNPFEDLYKAAELGIQVKNRNELEVAHRSIG 104
OY      169 TIOTATGLTERGIVLAPQIDKLLQK-TKAGQALGSAE-SIVONANKAKTVLSIGSIIG 226
Db      105 TADQFLGLTERGLTLFAPQIDQFLQKHSKISNVYSGTGDVASKLAKSQTIISGQSVLG 164
OY      227 SYLAGMDLEALONNSQNALAKAGLETNSLIENIANSVKTDERGEQISQPSKLONI 286
Db      165 TVLAGINLNEALITISGSELELAENGVSLSAEVSNIAKCTTTIDAFQTIONFGKLVENA 224
OY      287 KGLTGLDGLKNIIG--LDKAGGLDVLISGLSGATPAVLADKNASTAKKVGAGELAN 344
Db      225 KGLGAGVQROQNTISGSLSTKGLDITISLSGVASFALANKMNSTKVAAGFELSN 284
OY      345 QVVGNTIKAVSYTLAORVAAGLSSTGPVAALLASTVSLAISPLARAGIDKRNHAKSLE 404
Db      285 QVIGGIRKAVSYTLAORLAAGLSTGCPAALAAISSLSLISPLARVADNENRSKEIG 344
OY      405 SYAREFKLLDYDGNLAEYQRTGTIDASVTAINPLAIIAGVSAAD----- 455
Db      345 EPAREFKLLDYDGNLSEYHENGITIDASTITISRLSIAAGTAAGALVGPITL 404
OY      456 -----LTFEK-----VKHNLVITNSKKEKVTIONMFE----- 483
Db      405 LVYGITGLISGLEFSKQPMLDHVAIKGNKIDEMEKKYKNFYENGYDARHKAFLDSF 464
OY      484 ---ADFAKEVPNKRATK-----DEKTEELIIONGERTSKQVDDLIANK-----NG 526
Db      465 SLSSFEKKQYETERAVLITQORWDEYIGELAG---ITK--GDKLISSGKAYVDYFQEG 517
OY      527 KITQ---DELSKVVDNVELLKHKNVTNS---LDKLISSVSAFTSSNDR 570
Db      518 KLEKKRDKDDSKYV--FDPTKGEIDISNQTSLTKKVTPL--LTPTETSR 564

RESULT 9
RT31_ACTPL STANDARD; PRT; 1049 AA.
ID      RT31_ACTPL STANDARD; PRT; 1049 AA.
AC      P55130;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DE      30-MAY-2000 (Rel. 39, Last annotation update)
DE      RTX-III TOXIN DETERMINANT A FROM SEROTYPE 2 (APX-IIIA) (CYTOLYSIN
DE      IIA) (CLY-IIIA).
GN      APXIIIA OR CLYIIIA OR RTX A OR PTXA.
OS      Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC      Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC      Actinobacillus.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-SEROTYPE 2;
RX      MEDLINE-93263992; PubMed-8494611;
RA      Chang Y.-F., Shi J., Ma D.-P., Shin S.J., Iain D.H.;

```

RT "Molecular analysis of the *Actinobacillus pleuropneumoniae* RTX
 RL toxin-III gene cluster.";
 RN DNA Cell Biol. 12:351-362(1993).
 RP (2)
 RC SEQUENCE OF 828-1049 FROM N.A.
 RX STRAIN-1536 / SEROTYPE 2;
 RA MEDLINE=95012630; PubMed=7927703;
 RA Jansen R., Briatore J., van Geel A.B.M., Kamp E.M., Gielkens A.L.J.,
 RA Smits M.A.;
 RT "Genetic map of the *Actinobacillus pleuropneumoniae* RTX-toxin (Apx)
 RT operons: characterization of the ApxIII operon.";
 RL Infect. Immun. 62:4411-4418(1994).
 CC -1- FUNCTION: DOES NOT HAVE HEMOLYTIC ACTIVITY BUT SHOWS A STRONG
 CC CYTOTOXICITY TOWARDS ALVEOLAR MACROPHAGES AND NEUTROPHILS.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
 CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
 CC ACTIVITY (BY SIMILARITY).
 CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE
 CC INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).
 CC -1- PFM: PALMITOYLATED BY APMIIC. THE TOXIN ONLY BECOMES ACTIVE WHEN
 CC MODIFIED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L12145; AAA32924.1; -;
 DR EMBL: X80056; CAB37652.1; ALT-SEQ.
 DR INTERPRO: IPR001343; -;
 DR PFAM: PF00353; hemolysinCbind; 2.
 DR PROSITE: PS00330; HEMOLYSIN_CACIUM; 3.
 KW Toxin; Cytolysis; Cytotoxin; Repeat; Calcium; Transmembrane;
 KW Lipoprotein; Palmitate.
 FT TRANSMEM 154 170 POTENTIAL.
 FT TRANSMEM 315 331 POTENTIAL.
 FT TRANSMEM 397 413 POTENTIAL.
 FT DOMAIN 753 858 11 X REPEATS, GLY-RICH.
 FT REPEAT 753 758 1.
 FT REPEAT 762 767 2.
 FT REPEAT 771 776 3.
 FT REPEAT 780 785 4.
 FT REPEAT 789 794 5.
 FT REPEAT 798 803 6.
 FT REPEAT 807 812 7.
 FT REPEAT 826 831 8.
 FT REPEAT 835 840 9.
 FT REPEAT 844 849 10.
 FT REPEAT 853 858 11.
 SQ SEQUENCE 1049 AA; 112491 MW; F99846BFD4E5CE72 CRC64;

Query Match 26.1%; Score 938.5; DB 1; Length 1049;
 Best Local Similarity 43.5%; Pred. NO. 1.5e-39;
 Matches 231; Conservative 80; Mismatches 137; Indels 83; Gaps 16;

QY 109 GSPFKTGAKKIIYIPONYOYDFGNGLODLYKAAEELGVEYRENNRNATQTSIG 168
 DB 48 GRAVQKYG-NKLIVLYTPK--EYDGSVGNFGFDLYKAAEELGIVKYYNNRNELEVAHKSILG 104
 QY 169 TITQALGTERGIYLSAPQIDKLQK-TRQAQALGSAR-STYONANKAKTYLVSIGTISIG 226
 DB 105 TADQGLGTERGLTFLAPQLODFLOKHSKISNVGSGTDGASKAKSQTISIGTISIG 164
 QY 227 SYLAGMDLDELQONNSNOHAKKAGLELTNSLTENIANSVKTLDEFGQISQFSGSKLONI 286
 DB 165 TYLAGININLEAIISGSELELAAGVSLASLAVSNIAKGTITIDAFQTIONFGKLAENA 224

QY 287 KGLGTGDKLNKNGS--LDKAGLGLDVISGLSCATAALVLADKNASTAKKVGAGFELAN 344
 DB 225 KGLGCVGRQLOINISSALSKTGIGLIDISSLSGVTSTRFALRNKNASTSTVAKFELS 284
 QY 345 QVGNITKAVSSYIIAORVANGLSSTGVAALIASTVSLAISPAFACIADFNHAKSLE 404
 DB 285 QVIGITKAVSSYIIAORVANGLSSTGVAALIASTVSLAISPAFACIADFNHAKSLE 344
 QY 405 SYAEFRKRLGVDGDLAEYGRGTIDASTATATFALAAAGVSAAD----- 455
 DB 345 EFAEFRKRLGVDGDLAEYGRGTIDASTATATFALAAAGVSAAD----- 455
 QY 456 -----LPEEK-----YKHNVLITNSKKREKVTIONMFER----- 483
 DB 405 LVTGTLGSLGTERGPMMDHVAASKIGNKIDEMKRYGKNYPENGIDARKKAFLEDSF 464
 QY 484 ---ADFAKEVPNYATK-----DEKIEIIGONGERTITSKQVDDLAK-----NG 526
 DB 465 SLSSFNKQYETERAVLITQGRWDEYIGELAG-----ITGK--GDKLSSGRAVYDFQEG 517
 QY 527 KITQ---DELSKYVDNYELLHKSRYVTS---LDKLSSVSAFTSSDSR 570
 DB 518 KLEKKRPDFFSKVY--FDPTRGEIDISNSQSTSLKFTVPL--LTPGTESR 564

RESULT 10
 RT11.ACTPL STANDARD; PRT: 1023 AA.
 AC P55128;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE RTX-I TOXIN DETERMINANT A FROM SEROTYPES 1/9 (APX-IA) (HEMOLYSIN IA)
 DE (HLY-IA) (CYTOLYSIN IA) (CLY-IA).
 GN APXIA OR CLYIA OR HLYIA.
 OS *Actinobacillus pleuropneumoniae* (Haemophilus pleuropneumoniae).
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC *Actinobacillus*.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S 4074 / SEROTYPE 1;
 RX MEDLINE=91348845; PubMed=1879928;
 RA Frey J., Meier R., Gyg D., Nicolet J.;
 RT "Nucleotide sequence of the hemolysin I gene from *Actinobacillus*
 RT *pleuropneumoniae*.";
 RL Infect. Immun. 59:3026-3032(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S 4074 / SEROTYPE 1;
 RX MEDLINE=94237497; PubMed=8181764;
 RA Frey J., Haldemann A., Nicolet J., Boffini A., Prentki P.;
 RT "Sequence analysis and transcription of the apxi operon (hemolysin I)
 RT from *Actinobacillus pleuropneumoniae*.";
 RL Gene 142:97-102(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE CVI 13261 / SEROTYPE 9;
 RX MEDLINE=93366425; PubMed=8359891;
 RA Jansen R., Briatore J., Kamp E.M., Gielkens A.L.J., Smits M.A.;
 RT "Structural analysis of the *Actinobacillus pleuropneumoniae* RTX-toxin
 RT I (ApxI) operon.";
 RL Infect. Immun. 61:3688-3695(1993).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S 4074 / SEROTYPE 1;
 RX Chang Y., Wang Y., Chin N.;
 RL submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: ONE OF THE VIRULENCE FACTORS OF A PLEUROPNEUMONIAE,
 CC WHICH HAS A STRONG HEMOLYTIC ACTIVITY AND IS CYTOTOXIC FOR
 CC ALVEOLAR MACROPHAGES AND NEUTROPHILS.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
 CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC

```

CC          ACTIVITY.
CC          -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE
CC          INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).
CC          -1- PPM: PALMITOYLATED BY APXIC. THE TOXIN ONLY BECOMES ACTIVE WHEN
CC          MODIFIED (BY SIMILARITY).
CC          -1- MISCELLANEOUS: APXIA IS PARTIALLY DELETED IN SEROTYPES 2, 4, 6, 7,
CC          8, 12, AND TOTALLY DELETED IN SEROTYPE 3.
CC          -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF SEROTYPE 1.
CC          -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
CC          -----
CC          THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC          between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC          the European Bioinformatics Institute. There are no restrictions on its
CC          use by non-profit institutions as long as its content is in no way
CC          modified and this statement is not removed. Usage by and for commercial
CC          entities requires a license agreement (See http://www.isb.ch/announce/
CC          or send an email to license@isb.stb.ch).
CC          -----
CC          EMBL: X52899; CAA37081.1; -
CC          EMBL: X68595; CAA48586.1; -
CC          EMBL: X73117; CAA51548.1; -
CC          EMBL: U05042; AAB05034.1; -
CC          INTERPRO: IPR001343; -
CC          PPM: PPO0353; hemolysin_cabind; 2.
CC          PRINTS: PPO0313; CABDNGRPT.
CC          PROSITE: PS00330; HEMOLYSIN_CALCITUM; 2.
CC          Hemolysins; Toxin; Cytolysis; Cytotoxin; Repeat; Calcium;
CC          Transmembrane; Lipoprotein; Palmitate.
CC          TRANSMEM 226 256
CC          TRANSMEM 297 326
CC          TRANSMEM 367 406
CC          DOMAIN 722 845
CC          REPEAT 722 727
CC          REPEAT 731 736
CC          REPEAT 740 745
CC          REPEAT 749 754
CC          REPEAT 758 763
CC          REPEAT 767 772
CC          REPEAT 776 781
CC          REPEAT 785 790
CC          REPEAT 794 799
CC          REPEAT 813 818
CC          REPEAT 822 827
CC          REPEAT 831 836
CC          REPEAT 840 845
CC          REPEAT 210 217
CC          REPEAT 217 374
CC          REPEAT 374 562
CC          REPEAT 562 687
CC          REPEAT 687 1023
CC          SEQUENCE 1023 AA; 110193 MW; P99A88CFC9F1A598 CRC64;

Query Match      23.1%; Score 828.5; DB 1; Length 1023;
Best Local Similarity 37.3%; Pred. No. 4.3e-34;
Matches 216; Conservative 92; Mismatches 184; Indels 87; Gaps 19;

```

```

DB          247 VVASASFLSKNDADAGTAKAAGIEISTKILGNIGKAVSOYIIIAORVAAGLSTTAATGILI
DB          378 ASVVAISLPLAAGIADFNHAKSLSEYAEKPKKGYGDNILAYOGTGTIDSVTA 439
DB          307 GSVAVALAISPLSLNVADEFERAKOLEYSEKPKKGYEDDSILASFYETGATEALTT 366
DB          438 INTALAAIGVSAANA-----DLPFKVKNLVITNS 470
DB          367 INSVLARSAGVAAATGSLVGAFAVALVSAITGIIISILDASKAIFERVATKLANLKD 426
DB          471 KKEVITQWFE-----READFAKE---VPNYKATKDEKIEITIGONCE----- 511
DB          427 EWEKKGKKNFYENGVDARHSAPLEDFEELISQY--NKEYSVERVAVITQQRMDVINIGELA 484
DB          512 -ITSKQVDDILAG-----NGKITQDEL-----SKYVDNVELLKHKNYTNISLDR--LIS 558
DB          485 GILRKGSIDTKSGKAAVDFPEEGKLEKEDRDKRVDPLE-----GKIDLSINTKTYLLK 540
DB          559 SVS-AFTSSNDSRNLVAPTSMIDQSLSTLOFARGSQHW 596
DB          541 FVTPVFTAGEIRE---RKQTGKYQIMTEL-FVKGKEKW 575

RESULT 11
ID RT12.ACPL STANDARD; PRT: 1023 AA.
AC P55129;
DT 01-OCT-1996 (rel. 34, Created)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 30-MAY-2000 (rel. 39, Last annotation update)
DE RTX-1 TOXIN DETERMINANT A FROM SEROTYPES 5/10 (APX-1A) (HEMOLYSIN 1A)
DE (HLV-1A) (CYTOLYSIN 1A) (CLY-1A).
GN APXIA OR CLYIA OR HLXIA
OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Actinobacillus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13039 / SEROTYPE 10;
RX MEDLINE=94276858; PubMed=8007793;
RA Nagai S., Yagihashi T., Ishihama A.;
RT "DNA sequence analysis of an allelic variant of the Actinobacillus
RT pleuropneumoniae-RTX-toxin I (ApXIA) from serotype 10."
RL Microb. Pathog. 15:485-495(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K17 / SEROTYPE 5;
RX MEDLINE=96401417; PubMed=8807793;
RA Chin N., Frey J., Chang C.F., Chang Y.F.;
RT "Identification of a locus involved in the utilization of iron by
RT Actinobacillus pleuropneumoniae."
RL FEMS Microbiol. Lett. 143:1-6(1996).
RN [3]
RP SEQUENCE OF 886-1023 FROM N.A.
RC STRAIN=K17 / SEROTYPE 5;
RX MEDLINE=93366425; PubMed=8359891;
RA Jansen R., Briatore J., Kamp E.M., Gielkens A.L.J., Smits M.A.;
RT "Structural analysis of the Actinobacillus pleuropneumoniae-RTX-toxin
RT I (ApX1) operon."
RT Infect. Immun. 61:3688-3695(1993).
CC -1- FUNCTION: ONE OF THE VIRULENCE FACTORS OF A. PLEUROPNEUMONIAE,
CC WHICH HAS A STRONG HEMOLYTIC ACTIVITY AND IS CYTOTOXIC FOR
CC ALVEOLAR MACROPHAGES AND NEUTROPHILS.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE
CC INVOLVED IN PORE FORMATION BY THE CYTOTOXIN (BY SIMILARITY).
CC -1- PPM: PALMITOYLATED BY APXIC. THE TOXIN ONLY BECOMES ACTIVE WHEN
CC MODIFIED (BY SIMILARITY).
CC -1- MISCELLANEOUS: APXIA IS PARTIALLY DELETED IN SEROTYPES 2, 4, 6, 7,

```

8, 12, AND TOTALLY DELETED IN SEROTYPE 3.
 CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF SEROTYPE 10.
 CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D16582; BAA04014.1; -
 DR EMBL: U04954; AAB17220.1; -
 DR EMBL: X73116; CAA51546.1; -
 DR INTERPRO: IPR001343; -
 DR PIRAM; PF00353; hemolysinCbind; 2.
 DR PRINTS; PR00313; CABINDNGRPT.
 DR PROSITE; PS00330; HEMOLYSIN_CALCIUM; 2.
 KW Hemolysin; Toxin; Cytolysin; Cytotoxin; Repeat; Calcium;
 KW Transmembrane; Lipoprotein; Palmitate.
 FT TRANSMEM 226 256
 FT TRANSMEM 297 326
 FT TRANSMEM 367 406
 FT DOMAIN 722 845
 FT REPEAT 722 727
 FT REPEAT 731 736
 FT REPEAT 740 745
 FT REPEAT 749 754
 FT REPEAT 758 763
 FT REPEAT 767 772
 FT REPEAT 776 781
 FT REPEAT 785 790
 FT REPEAT 794 799
 FT REPEAT 813 818
 FT REPEAT 822 827
 FT REPEAT 831 836
 FT REPEAT 840 845
 FT REPEAT 840 845
 FT CONFLICT 210 217
 FT CONFLICT 581 581
 FT CONFLICT 687 688
 FT CONFLICT 1015 1015
 SQ SEQUENCE 1023 AA; 110129 MW; 183C7C15EE57DB55 CRC64;
 Query Match 23.1%; Score 828.5; DB 1; Length 1023;
 Best Local Similarity 36.7%; Pred. No. 4.3e-34;
 Matches 213; Conservative 91; Mismatches 187; Indels 89; Gaps 17;
 Oy 85 SQHMEYGLRPGSGSDMSYGLRPGSSFPKTKGAKKIIILYIPNOYVDIEOGGLDDLYKA 144
 Db 17 NQHTSKAASGGA-----LNKNGQYKQAG-QKILYIPDYAST--GSLNDLYKA 67
 Oy 145 AEELIEVQREERNIATAQTSLGTQTAIGTERGIYLSAPQIDKLQKT-KAGQAL-G 202
 Db 68 AEALGIEVHRSKKNTALAKELFGTEKILGSEKGIALFAQFPDKLNKNKLSLGG 127
 Oy 203 SAESIVQANKAKIVLSGIQSLISVLGMDLDEALQNNNSNH-----ALARAQELTNS 257
 Db 128 SEALGQRILKQTLALNALQSFALGAGMDLSLRRRRNCEDEVSGSLAKAGVDLAAQ 187
 Oy 258 LIENIANSVKTLDEGEQDISQSGSKLQNKIGTIGDKRIKNGIGDKAGLGDIVISGLIS 317
 Db 188 LVDNIAASAGIVDAFEDQGLKAMPYLYL-ALSGLASKILNNL-PDLISLGGPGFDAVSGILS 246
 Oy 318 GATALLVADKNASTAKKVGAGFELANQVGNITKAVSSYIIAQRVAAGLSTGPPAALI 377
 Db 247 VVSASFILSNKQADAGTAAAGIEISTKILNIGKAVSYIIIAQRYAAGLSTTAATGILI 306
 Oy 378 ASTVSLAIPLAFAIGADKNFNHAKLSLEYAEFRKKLGYDGNLLAEYQGTITDASVTA 437
 Db 307 GSVVALAISPLSFLNVADFERAKQLEQYSERFKKFGYEGDLSLASFYRETGAIEALIT 366

Oy 438 INTALAAAGVSAAAA-----DLFEKYHNIVTNS 470
 Db 367 INSVALSASAGGAATGSLVAPYALVSATITGIISGLDASKAITEFVATKLANMD 426
 Oy 471 KKEVLTIONMF-----READFAKE---VPNYKATKDEKIEIIGONGER----- 511
 Db 427 EWEKHKKNYFENGVDARHSAPLEDFTELLSQY--NKEYSVRYVAIIQOQPDVINGELA 484
 Oy 512 -ITSQVDDLLKGG-----NKRITODELSKYVDYELLKHSKNYNTNSDKLI--SSVSA 562
 Db 485 GTRRGAARSKAGKAVYDFEEEGKL---LEKPDREFD-----KKVDFPLEGRIDLSINK 535
 Oy 563 FTSSNDSRVLYAPTSMDQSLSSLO-----FAGSGOHM 596
 Db 536 TTLKFTTPVFTAGEEIERKOTGKYEWTELPVAGKEW 575
 RESULT 12
 ID HLY1_ECOLI STANDARD: PRT: 1023 AA.
 AC P09983;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE HEMOLYSIN, CHROMOSOMAL.
 GN HLYA.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-196 / SEROTYPE O4;
 RX MEDLINE-85234404; PubMed-3891743;
 RA Felmlee T., Pellett S., Welch R.A.;
 RT "Nucleotide sequence of an Escherichia coli chromosomal hemolysin";
 RL J. Bacteriol. 163:94-105(1985).
 RN (2)
 RP SEQUENCE OF 1-44 FROM N.A.
 RC STRAIN-2001;
 RX MEDLINE-85258115; PubMed-3894051;
 RA Nicaud J.-M., Mackman N., Gray L., Holland I.B.;
 RT "Characterisation of HlyC and mechanism of activation and secretion
 of haemolysin from E. coli 2001";
 RL FEBS Lett. 187:339-344(1985).
 CC -1- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
 CC CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
 CC DEFINED.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
 CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
 CC ACTIVITY.
 CC -1- DOMAIN: THE THREE TRANSMEMBRANE DOMAINS ARE BELIEVED TO BE
 CC INVOLVED IN PORE FORMATION BY THE CYTOTOXIN.
 CC -1- PTM: PALMITOYLATED BY HLYC. THE TOXIN ONLY BECOMES ACTIVE WHEN
 CC MODIFIED.
 CC -1- DISEASE: THE HEMOLYSIN OF E. COLI IS PRODUCED PREDOMINANTLY BY
 CC STRAINS CAUSING EXTRAINTESTINAL INFECTIONS, SUCH AS THOSE OF THE
 CC URINARY TRACT.
 CC -1- SIMILARITY: BELONGS TO THE RTX PROKARYOTIC TOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M10133; AAA23975.1; -
 DR EMBL: X02768; CAA26546.1; -
 DR PIR: A24433; LEECA.
 DR INTERPRO: IPR001343; -
 DR PIRAM; PF00353; hemolysinCbind; 2.


```

DR INTERPRO: IPR001343;
DR PFAM: PF00353; hemolysinCbind; 5.
DR PRINTS: PR00313; CABINDNGRPT.
DR PROSITE: PS00330; HEMOLYSIN_CALCTIM; 5.
KW Lyase; CAMP synthesis; ATP-binding; Hemolysis; Toxin; Virulence;
KW Whooping cough; Calcium-binding; Repeat; Lipoprotein; Palmitate.
FT CHAIN 1 312 CALMODULIN-SENSITIVE ADENYLATE
FT CHAIN 1 312 CYCLASE.
FT CHAIN 313 1706 HEMOLYSIN (BY SIMILARITY TO E.COLI
FT 1 HEMOLYSIN HLVA).
FT DOMAIN 1 399 A, CATALYTIC.
FT DOMAIN 400 912 B, ALA/GLY-RICH.
FT DOMAIN 913 1656 C.
FT DOMAIN 1657 1706 D. ASP/GLY-RICH.
FT NP_BIND 349 356 ATP (POTENTIAL).
FT DOMAIN 913 1610 28 X REPEATS, GLY-RICH.
FT REPEAT 1015 1020 1.
FT REPEAT 913 918 2.
FT REPEAT 1024 1029 3.
FT REPEAT 1033 1038 3.
FT REPEAT 1042 1047 4.
FT REPEAT 1051 1056 5.
FT REPEAT 1060 1065 6.
FT REPEAT 1080 1085 7.
FT REPEAT 1165 1170 8.
FT REPEAT 1174 1179 9.
FT REPEAT 1183 1188 10.
FT REPEAT 1203 1208 11.
FT REPEAT 1280 1285 12.
FT REPEAT 1289 1294 13.
FT REPEAT 1298 1303 14.
FT REPEAT 1307 1312 15.
FT REPEAT 1316 1321 16.
FT REPEAT 1325 1330 17.
FT REPEAT 1345 1350 18.
FT REPEAT 1421 1426 19.
FT REPEAT 1430 1435 20.
FT REPEAT 1439 1444 21.
FT REPEAT 1448 1453 22.
FT REPEAT 1456 1461 23.
FT REPEAT 1565 1570 24.
FT REPEAT 1574 1579 25.
FT REPEAT 1583 1588 26.
FT REPEAT 1593 1598 27.
FT REPEAT 1605 1610 28.
FT LIPID 860 860 PALMITATE.
FT LIPID 983 983 PALMITATE.
FT MUTAGEN 188 188 D->E,N,Y,H: LOSS OF ACTIVITY.
FT MUTAGEN 190 190 D->N,Y,H: LOSS OF ACTIVITY.
FT MUTAGEN 298 298 H->R,P,L: LOSS OF ACTIVITY.
FT MUTAGEN 301 301 E->Q,K: LOSS OF ACTIVITY.
SQ SEQUENCE 1706 AA; 177506 MW; F00744524BDD442E CRC64;

```

Query Match 5.9%; Score 212; DB 1; Length 1706;

Best Local Similarity 22.3%; Pred. No. 0.003; Indels 94; Gaps 21;

Matches 114; Conservative 84; Mismatches 220; Indels 94; Gaps 21;

```

QY 89 SYGLRPGSGSODWSYGLRPGSGSPFKTAKKIIILYIPONYQYDTQNGIOLDLVRAAEEL 148
DB 349 AYGV-AGKSLFDGAGAPGVSPGSKSPDLVETVPAS-----PGLR-----RPSL 394
QY 149 GIEVOREER-----NNIATAOTSIGTIOAIGLTERGIVLSAPQIDKLLQTKKAGQALGS 203
DB 395 G-AVERODSGYSLDGVSGSRFSLEGVSDM-----AAVEAAELEMTROYVLHAGARQDD 446
QY 204 AESTIVQNK--AKTVLGGIOSILGSVLAGMDLEALQNNNOHALA-----KAGLELT 255
DB 447 AEPGVSGASAHMGORALQAOQAVANA-----QRLVHAIALMTOPGRAGSTNT 493
QY 256 NSLIENIANSVTLDFEFGQISQFGSKLQINIGLGLDKLKNIGLKDAGLGLDVISGL 315
DB 494 PQEASLSAAVVGCLGEASAAVAETVS-----GFFRSSRWAGGFGVAGGAMALGGCI 545

```

```

QY 316 LSGATAALVLDKKNSTAKKVGAGFELANOVYGNITKAVSYIL-----AQRYAAGLSSTG 371
DB 546 AAAGVAGMSLTD-DAPAGQKAAAGAEIALQLTGTVELASSITALALAAARGVTSGLQVAG 604
QY 372 PYAALIASTVSLAISPLAFAGIADKFNHAKSLESYAERPKKLGYGDNLLAEYQGTGTI 431
DB 605 ASAGAAAGALAAALPMEIYGLVQOSHADQDLKLAQESSAAGYEGDALIAQLYRDKTAA 664
QY 432 DASVTAINATALAAGVVS-AAADLTFEKYNHNLVTNSKKEKVT-----IQQNFREA- 484
DB 665 ECAVAGVSALSTVGAAVSIAAASVGAHV---AVVTSLLTGALNGILRGVQOPFIEKL 721
QY 485 --DFAKEVPN-----YKATDEKIEELIGNGERITRSKOYDDLIANGNR-----ITQD 531
DB 722 ANDYARKIDELGGPOAYFEKNLQARHEQLANSDDLRL---KMLADLQAGNMASVIGVQTT 778
QY 532 ELKVVVDNYELLKHSKNVTNSLDKLISSVSAP 563
DB 779 EISK--SALELAATTGNADN-----LKSVDVF 803

```

Search completed: March 2, 2001, 11:01:34
Job time: 449 sec

THIS PAGE BLANK (USPTO)

OM of: US-09-306-689-12 to: SPTREMBL_15:* out_format : pfs

Date: Mar 2, 2001 10:49 AM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frmet+npz.model -DEV=rlp
-Q/cgn2.1/USF70.spool/US09306689/runat.02032001_102828_9616/app_query.fasta.1.2389
-DB=SPTREMBL_15 -QFMT=fasta -SUFFIX=rsp -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
-GAPOP=4.000 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FCAPOP=6.000 -FCAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELDP=6.000 -DELEXT=7.000 -START=1 -MATRIX=bloms62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09306689_@CGN1_1_639 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPHY
-MIT -THREADS=1

Search information block:

Query: US-09-306-689-12

Query length: 2088

Database: SPTREMBL_15:*

Database sequences: 374700

Search time (sec): 337.480000

score_list:

Sequence	Strd	Orig	ZScore	EScore	len	Documentation
SP_bacteria:068403	+	377.50	366.99	1.2e-12	181	068403 escherichia coli. alpha
SP_bacteria:068404	+	377.50	366.99	1.2e-12	181	068404 escherichia coli. alpha
SP_bacteria:070070	+	372.50	364.28	2.3e-12	181	070070 escherichia coli. alpha
SP_invertebrate:09617	+	208.50	191.73	0.0010	1729	09617 drosophila melanogaster
SP_bacteria:091469	+	205.50	189.00	0.0014	1706	091469 bordetella parapertussis
SP_bacteria:098433	+	199.00	177.44	0.0031	3381	098433 streptococcus cristatus
SP_bacteria:052781	+	190.50	178.27	0.0085	1112	052781 campylobacter fetus. su
SP_bacteria:061889	+	187.50	179.04	0.0121	707	061889 mus musculus (mouse). ke
SP_bacteria:091964	+	187.50	176.93	0.0121	922	091964 mus musculus (mouse). ke
SP_invertebrate:09766	+	185.50	174.09	0.0155	1039	09766 drosophila melanogaster
SP_invertebrate:013083	+	173.00	165.20	0.0697	722	013083 dissostichus mawsoni. at
SP_bacteria:098073	+	170.00	164.32	0.0997	565	098073 escherichia coli. flagel
SP_bacteria:098473	+	169.00	155.79	0.1147	1467	098473 deinetococcus radiodurans
SP_bacteria:098473	+	168.50	156.99	0.1213	1190	098473 xylella fastidiosa. su
SP_bacteria:068729	+	166.50	153.13	0.1553	1525	068729 yersinia pestis. prote
SP_bacteria:098473	+	166.00	158.24	0.1627	756	098473 deinetococcus radiodurans
SP_bacteria:053505	+	165.00	154.25	0.1850	1109	053505 campylobacter fetus. ty
SP_invertebrate:097054	+	165.00	147.43	0.1882	2614	097054 dictyostellum disco
SP_bacteria:047226	+	164.00	156.67	0.2060	565	047226 escherichia coli. flagel
SP_bacteria:080179	+	163.00	149.91	0.2371	1510	080179 streptococcus thermoph
SP_invertebrate:098473	+	162.00	150.02	0.2668	1323	098473 deinetococcus radiodurans
SP_fungi:013936	+	161.00	152.25	0.2994	990	013936 schizosaccharomyces pom
SP_invertebrate:098473	+	160.50	151.38	0.3170	836	098473 deinetococcus radiodurans
SP_bacteria:098473	+	160.00	144.61	0.3428	2059	098473 xylella fastidiosa. su
SP_bacteria:034071	+	159.00	146.10	0.3845	1517	034071 streptococcus thermoph
SP_invertebrate:098473	+	159.00	146.10	0.3845	1517	098473 deinetococcus radiodurans
SP_bacteria:046037	+	158.50	149.47	0.4046	936	046037 campylobacter fetus. sat
SP_bacteria:098473	+	158.50	144.30	0.4099	1794	098473 deinetococcus radiodurans
SP_bacteria:098473	+	158.00	147.32	0.4316	1156	098473 deinetococcus radiodurans
SP_bacteria:098473	+	158.00	146.22	0.4328	1328	098473 deinetococcus radiodurans
SP_bacteria:054415	+	157.50	153.40	0.4511	507	054415 salmonella montevideo. ph
SP_bacteria:097643	+	157.00	153.67	0.4784	462	097643 lama glama (llama). fib
SP_bacteria:098473	+	156.00	147.49	0.5469	892	098473 deinetococcus radiodurans
SP_bacteria:054515	+	155.50	151.50	0.5746	508	054515 salmonella pensacola. ph
SP_bacteria:098473	+	155.50	145.59	0.5831	1068	098473 deinetococcus radiodurans

seq_name: sp_bacteria:068403

seq_documentation_block:

ID 068403 PRELIMINARY; PRT: 181 AA.

AC 068403:
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE ALPHA HEMOLYSIN (FRAGMENT).
GN HYL.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
NCBI_TaxID=562;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ECOR 24;
RX MEDLINE=98155142; PubMed=9495754;
RA Boyd E.F., Hartl D.L.;
RT "Chromosomal regions specific to pathogenic isolates of Escherichia coli have a phylogenetically clustered distribution".
RL J. Bacteriol. 180:1159-1165(1998).
DR EMBL: AF037572; AAC38261.1;
DR INTERPRO: IPR000911;
DR PROSITE: PS00359; RIBOSOMAL_L11; UNKNOWN_1.
FT NON_TER 1
FT SEQUENCE 181 AA; 19071 MW; 570928D2233B4C04 CRC64;

alignment_scores:
Quality: 377.50 Length: 181
Ratio: 2.640 Gaps: 3
Percent similarity: 79.006 Percent identity: 45.856

alignment_block:

US-09-306-689-12 x 068403 ..

Align seg 1/1 to: 068403 from: 1 to: 181

403 GGTAATGCTTACAGATTATGACCAAGCGCGCAAGTGGATTTGA 452
1 GlySerSerLeuAsnSpleuValArgThrAlaAspGluLeuGlyIleG 17
453 GGTCAAGAGAGAGAGCAATATATTCACACAGCTCAACCCATTAG 502
17 ValGlnTyrAspGluTyrAsnGlyThrAlaThrIleThrValAla 34
503 GCAAGATTCAAAACCGTATTGCTTACAGAGCGTGGCATTTGTTATCC 552
34 LThrAlaGluTyrLeuIleGlyLeuThrGluArgGlyValAlaIle 50
553 GCTCAAAATGATTAATTTGATCAGAAA...ACTAAAGAGCGCAAGC 599
51 AlaProGlnLeuAspTyrLeuGlnTyrGlnIleValAlaGlyAsn 67
600 ATTA...GGTTGCGCAAGCAATTTGATCAAAATTAAGCAAAA 646
67 sleuGlyGlySerAlaGluAsnIleGlyAspAsnIleGlyValAla 84
647 CAGTATTATTCGATCAATCAATTTATTTAGCGTCAGTATTTGCTGA 696
84 ervaLeuSerThrPheGlnAsnIlePheLeuGlyAlaAlaLeuSerMet 100
697 GATTATTCAGGCTTCACAG.....ATAACAGCAACCA 731
101 LysIleAspGluLeuIleLysGlnIleLysSerGlyAsnValSer 117
732 ACATGCTTGTCTTAAGCTGCTGGAGCTTAACAAATTCATTAATGAA 781
117 rSerGluLeuIleValSerIleGluLeuIleAsnGlnLeuValAsp 134
782 ATATTGCTTAATTCAGTAACAAATTCAGCAATTTGCTGACCAATTA 831
134 hValAlaSerIleAsnAsnValAsnSerPheSerIleGlnLeuAsn 150
832 CAATTGCTTCAAACTCAAAATATCAAGCTTATGAGCACTTATGACA 881


```

828 TAGTCATTTGGTTCAAACCTACAAATATCAAGGCTTAGGACT... 873
    |||::: ||| |||||
693 oGlyIleGlyGlnThrGlyProGlyGlnProGlyTyrGlyThrGlnP 710
    |||::: ||| |||||
874 .....TTAGAGACAAACTC..... 888
    |||::: ||| |||||
710 roGlyValGlyThrGlnThrGlyThrGlyGlnProGlyTyrGlyAlaGln 726
    |||::: ||| |||||
889 AAAAATTCGGTGGCTGATTAAGCTGGCCTT..... 921
    |||::: ||| |||||
727 ProGlyIleGlyGlnSerGlyAlaGlyGlnProGlyTyrGlySerGly 743
    |||::: ||| |||||
922 .....GGTTAGATGTTATCTCAGGAC 943
    |||::: ||| |||||
743 nProGlyIleGlyGlnThrGlyGlnProGlyTyrGlySerGly 760
    |||::: ||| |||||
944 TATTATCGGGCCACAGCTGCACCTTACTTGCAGATTAATAATGCTTCA 993
    |||::: ||| |||||
760 InIleGlyGlyGlnThrGlyAla.....GlyGlnProSer 771
    |||::: ||| |||||
994 ACAGCTAAAGTGGGGGTTTGAATTTGGCAAAACCAAGTTGTTGG 1043
    |||::: ||| |||||
772 TyrGlySerGlnProGlyValGlyAlaGlnAsnGlyGlyGlnProGly 788
    |||::: ||| |||||
1044 TAAATATCCAAAGCCGTT..... 1062
    |||::: ||| |||||
788 YTYrGlyTYrThrArgProValIleGlyGlyGlnThrGlyAlaGlyGlnProG 805
    |||::: ||| |||||
1063 .....TCTTCTTACATTTTACCCCA 1083
    |||::: ||| |||||
805 LYTYrGlyGlyGlnThrGlyValGlyGlySerProGlyPheLeuThrGln 821
    |||::: ||| |||||
1084 CGTGTTCAGAGGTTATCTCAACTGGCGCTGGCTGCTGTTAAATTCG 1133
    |||::: ||| |||||
822 ProGlyIleGlyGlyIleSer.....GlyProIleGlyGlyValGly 836
    |||::: ||| |||||
1134 TTTACTGTTTCTCTGCAATTAGCCCATTAAGATTGCC..... 1173
    |||::: ||| |||||
836 yGlyGlyGlnSerGlnAlaAlaLysProGlyTYrTYrPalGlnProGlyI 853
    |||::: ||| |||||
1174 .....GGTATTGCCGATTAATTT 1191
    |||::: ||| |||||
853 leGlyGlyProSerArgTYrGlySerGlnProGlyIleGlyLysPheLthr 869
    |||::: ||| |||||
1192 AATCATCAAAAGTTAGAGATTATGCCAAGCTTTAAATAATTAG 1241
    |||::: ||| |||||
870 GlyAlaLeuGlnSer.....Gly 875
    |||::: ||| |||||
1242 CTATGACGAGATTAATTATTAGCAGATATCAGCGGGAGACGAGCTA 1291
    |||::: ||| |||||
875 YTYrGlyGlyGlnProGlyIleSerGly...GlnThrGlyGlyGlnP 891
    |||::: ||| |||||
1292 TT.....GATGCATCGGTACTGCAATTAAT..... 1317
    |||::: ||| |||||
891 roGlyTYrGlyGlyGlnAlaThrIleSerGlyLeuProGlyTYrGlyThr 907
    |||::: ||| |||||
1318 .....ACGCATTGGCCGCTATTGCTGCTGCTGCTGCTGCTGCTG 1355
    |||::: ||| |||||
908 GlnProGlyIleGlyAlaLeuThrAlaValProGlyGly..... 920
    |||::: ||| |||||
1356 TGCAGCCGATTTAACATTGAAAAAGTTAAACATTAATCTGTCTATCAG 1405
    |||::: ||| |||||
921 .....HistyGlyTYrGln..... 925
    |||::: ||| |||||
1406 ATACAAAAAGAGAAAGTGCACATTCAAAACGTGTTCCGAGGCTGAT 1455
    |||::: ||| |||||
926 .....ThrGlnProGlyIleGlyGlyGlnThrGlyThrAsnGlnProGly 940
    |||::: ||| |||||
1456 TTTGCTAAAGAAAGTGCCTAATTATTAAGCACTAAAGATGAGAAATGCA 1505
    |||::: ||| |||||
941 PheGlyGlyGln.....Pr 945
    |||::: ||| |||||
1506 AGAAATCATCGTCAAAATGCGGAGCGGATCACCTCAAGCAAGTGTGATG 1555
    |||::: ||| |||||

```

```

945 oGlyIleGlyGlnThrGlyAlaGly..... 954
    |||::: ||| |||||
1556 ATCTTATCGCAAAAGTAAGGCAAAATTAACCAAGATGAGCTATCAAAA 1605
    |||::: ||| |||||
955 .....GlnProGlyTYrGlyPheIleGlyGlnProGlyIleGlyGly 968
    |||::: ||| |||||
1606 GTTGTGATTAACATGAAATGCTCAAAACATAGCAAAATGTGACAAACAG 1655
    |||::: ||| |||||
968 ..... 968
    |||::: ||| |||||
1656 CTTAGATTAAGTTAATCTCATCTGTAAGTCAATTACTCGCTAATGATT 1705
    |||::: ||| |||||
968 ..... 968
    |||::: ||| |||||
1706 CGAATAATGATTAGTGGCTCCAACTTCATGTTGATCAAGTTTATCT 1755
    |||::: ||| |||||
968 ..... 968
    |||::: ||| |||||
1756 TCTCTCAATTTGCTTAGGAGATCTACAGATTTGAGCTACGCGCTGCGCC 1805
    |||::: ||| |||||
969 .....GlnThrGlyThrSerGlyArgGlnProGlyTYrGlyThrGlnP 983
    |||::: ||| |||||
1806 TGCAGCGGTTCTCAGATTGAGC.....TACGCGCTGCGCTC 1843
    |||::: ||| |||||
983 oGlyIleGlyGlyGlnThrAlaAlaGlyGlnProGlyTYrGlySerGlnP 1000
    |||::: ||| |||||
1844 CGGCT.....GGCTCTAGCCAGCATTTGAGCTACGCGCTG 1878
    |||::: ||| |||||
1000 roGlyIleGlyGlyGlnThrGlyAlaGlyGlnPro...GlyTYrGlySer 1015
    |||::: ||| |||||
1879 CGCCCTGCGAGCGGTAGCCAA.....GATTGAGCTACGCGCT 1916
    |||::: ||| |||||
1016 GlnThrGlyValGlyGlyGlnIleGlyAlaGlyGlnProGlyTYrGlyse 1032
    |||::: ||| |||||
1917 GCGTCCGGCT.....GGATCTCAGCATTTGAGCTACGCGCC 1951
    |||::: ||| |||||
1032 rGlnProGlyIleGlyGlyGlnThrGlyAlaGlyGlnProGlyTYrGly 1049
    |||::: ||| |||||
1952 TGGCGCCGCGAGCGGTTTCACAGATTGAGCTACGCGCTGCGCGGT 2001
    |||::: ||| |||||
1049 laGlnProGlyPheGlyGlyGlnPro...GlyTYrGlyAsnGlnProGly 1064
    |||::: ||| |||||
2002 .....GGCTTAGCCAGCATTTGAGCTACGCGCTGCGGCC 2036
    |||::: ||| |||||
1065 ValGlyGlyGlnThrGlyAlaGlyGlnPro...GlyTYrGlySerGlnP 1080
    |||::: ||| |||||
2037 TGCAGCGGTAAGCCAGATTGAGC.....TACGCGCTGCGCTC 2074
    |||::: ||| |||||
1080 oGlyValGlyGlyGlnThrGlyAlaGlyGlnProGlyTYrGlyValIleP 1097
    |||::: ||| |||||
2075 CGGCT 2079
    |||::: ||| |||||
1097 roGly 1098
    |||::: ||| |||||

```

seq_name: sp_bacteria:Q91A69

```

seq_documentation_block:
ID Q91A69 PRELIMINARY: PRT: 1706 AA.
AC Q91A69;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE BIFUNCTIONAL HEMOLYSIN-ADENYLATE CYCLASE PRECURSOR (Ec 4.6.1.1).
GN CYA.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=519;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=63.2;
RA Boursaux-Ende C., Guiso N.;

```

"Sequencing of the Bordetella parapertussis adenylate cyclase-hemolysin-encoding gene";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ249835; CAB76450.1; -
 KW Lyase.
 SQ SEQUENCE 1706 AA; 177038 MW; 71750E2D0BBF3B64 CRC64;

alignment_scores:

Quality: 205.50 Length: 477
 Ratio: 0.849 Gaps: 16
 Percent Similarity: 50.734 Percent Identity: 22.432

alignment_block:

US-09-306-689-12 x Q9L469 ..

Align seg 1/1 to: Q9L469 from: 1 to: 1706

```

16 GATCGATCTCAGATTCGAGCTACGCGCTCGCGCCCTGGCAGCGGCTTCGA 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
294 AspValValAlGlnHisGlyThrGlnAsnAspProPheProGlnAlaAs 310
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
66 AGATTGGAGCTACGCGCTCGCGGCTGCTTACCCAGCATTTGGAGCT 115
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
310 pGluLysIlePheValValSerAlaThrGlyLysSerGlnMetLeuThr 327
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
116 AGGCG...CTGCGCCCTGGCAGCGGTACCA... 144
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
327 rGlyGlnLeuLysGlnLysIleGlyGlnGlnArgLysGlnLysIleVal 343
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
145 .....GATTGAGCTACGCGCTCGCGGCTGATTCACAGATTCG 185
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
344 PheTyrGlnAsnArgAlaTyrGlyVal.....AlaGlyLysSerLeu 358
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
186 GAGCTACGCGCTCGCGCTCGCGGCTTCGAAGATTGGAGCTACGCGC 235
   :::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
358 eAspAspGlyLeuGlnAlaAlaProGlyVal..... 368
236 TCGCTCGCGGCTGCTTACCCAGCATTTGAGCTACGCGCTCGCGCTG 285
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
369 .....ProGlyLysArgSerLys..... 374
286 AGCGGTACGCGCAT.....TGGAGCTACGCGCTCGG 317
   :::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
375 ...SerSerProAspValLeuGlnThrValProAlaSerProGlyLeu 390
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
318 TCCGGGTGATCTAGCTCCCAAAACTGGGCAAAAAAATTATCTCT 367
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
390 gArg.....ProSerLeuGlnAlaValGlu..... 398
368 AATATCCCAAAATTACCAATATGATCTGAACAAGATTAAGTTTACAG 417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
399 .....ArgGlnAspSerGlyTyrAspSerLeuAspIleValGlySerArg 413
418 GATTTAGTCAAAAGCGGCCGAAGGTTGGGATTTAGGTACAAAGACAAA 467
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
414 SerPhe..... 415
468 ACGCAATATATTGCAACAGCTCAACAGATTAGCAGATTCACAAACG 517
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
416 .....SerLeuGlnLysValSerAspM 423
518 CTAATGGCTTAACGTAGCGCTGATTTCCGCTCCACAAATTTGAT 567
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
423 et.....AlaAlaValAlGlnAlaValGlnLeuGln 432
568 AATATGCTACAGAAAAGTAAGAGCGCAAGCATTTAGTTGGCCGAAG 617
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
433 MetThrArgGlnValLeuHisAlaGlyAlaArgGlnAspAspAlaLys 449
618 CATTTGACAAATATGCAATATAA.....GCCAAACTGATTATATCGGCA 661
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
449 oGlyValSerGlyAlaSerAlaHisTyrpGlyLysArgAlaLeuGlnGly 466

```

```

662 TTCATCTATTTTAGCTCAGATTGGCTGGAATGATTCATTACATGAGCC 711
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
466 lAglnAlaValAlaAlaAla..... 472
712 TTACAGATTAACGCAACCAACATGCTTGGT..... 744
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
473 .....GlnArgLeuValHisAlaIleAlaLeuMetThrGlnPheG 486
745 .AAGCTGGCTTGGAGCTTACCAAAATTCATTAAATGAAAATTTGCTAAT 793
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
486 yArgAlaGlySerThrAsnThrProGlnGlnAlaAlaSerLeuSerAla 503
794 CAGTAAAAACACTTGACGAATTTGGTGAACAAATTAAGTAATTTGGTTCA 843
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
503 lValAlaPheGlyLeuGlnGlyAlaAlaSerSerAlaValAlaGlnThrValSer 519
844 AAACCTACAAATATATCAAGCTTAGGACTTTAGAGCAACAACTCAAAA 893
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
520 .....GlyPhePheArgGlySerSerArgTr 528
894 TATCGGTGACTTGAATTAAGCTGGCTTGGTTAGATTTATCTCAGGCG 943
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
528 pAlaGlyGlyPheGlyValAlaGlyAlaMetAlaLeuGlyGlyI 545
944 TATTTATGGGCGCCACAGCTGCTTGTACTTGCAGATTAATAATGCTTCA 993
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
545 lAlaAlaAlaValAlaGlyAlaGlyMetSerLeuThrAsp...AspAlaPro 560
994 ACAGCTAAAAAAGTGGGCTGGGTTTGAATGGCAACCAAGTTGGTGG 1043
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
561 AlaGlyGlnLysAlaAlaValAlaGlyAlaGlnIleAlaLeuGlnLeuThr 577
1044 TATATTACCAAGCGCTTCTTCTTACATTTA.....GCCC 1081
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
577 yGlyThrValGlnLeuAlaSerSerIleAlaLeuAlaLeuAlaAla 594
1082 AACGTGTTCACAGAGTTTATCTCACTAGCGCTGTGGCTCTTAAAT 1131
   :::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
594 rGlyValThrSerGlyLeuGlnValAlaGlyAlaSerAlaGlyAlaAla 610
1132 GCTTCTACTGTTTCTTCTGCAATAGCCCATTAAGCATTTGCCGATTTGC 1181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
611 AlaGlyAlaLeuAlaAlaAlaLeuSerProMetGluIleTyrGlyLeuVal 627
1182 CGATAAATTTAATTCATGCAAAAGTTTAGAGATTAGCCGAACCTTTA 1231
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
627 lGlnGlnSerHisTyrAlaAspGlnLeuAspLysLeuAlaGlnGlnSer 644
1232 AAAAATTTAGCTATGACGAGATTAATTTATAGCAAGATATACGCGGGA 1281
   :::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
644 eAlaIleTyrGlyGlnGlyAlaAspAlaLeuAlaGlnLeuTyrArgAsp 660
1282 ACAGGAGCTATGATGATCGGTTACTGCAATTAATACCGATTGGCGCC 1331
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
661 LysThrAlaAlaGlnGlyAlaValAlaGlyValSerAlaValLeuSerThr 677
1332 TATTCGTGCTGTTGTTCTGCTGCTGACGCC 1362
   :::: :::: :::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
677 rValGlyAlaAlaValSerIleAlaAlaAla 687

```

seq_name: sp_bacteria:Q9KX33

seq_documentation_block:

ID Q9KX33 PRELIMINARY; PRT; 3381 AA.

AC Q9KX33:

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

DE SRPA.

OS Streptococcus cristatus.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.


```

1497 GAAAAATCGAAGAAATCATCGTCAAAATGGCGAGCGGATCACCCTCAAGC 1546
      :::::::::::::::::::: |||::: :::: |||||
2751 nSerValSerIleSerIleSerIleSerValSerIleSerThrSer.... 2766
      |||
1547 AAGTTGATGATCTTATCCGAAAAGTACGCAAAATTTACCAAGATGAG 1596
      ::::::::::::::::::::
2767 .....MetSerAsnSer 2770
1597 CATCAAAAGTTGTGATPACTATGTAATTCGCAAAACATGCAAAATAT 1646
      :::::::::::::::::::: |||
2771 ValSerIleSerIleSerIleSerIleSerIleSerIleSerIleSerIle 2785
      |||
1647 GACAACAGCTTAGATTAAGTTATCTCA...TCTGTAAGTGCATTATACCT 1693
      :::::::::::::::::::: |||
2785 tSerIleSerValSerIleSerIleSerIleSerIleSerIleSerIleSer 2802
      |||
1694 CCGTAAATGATTCGAGAAATGATTATAGTGGTCCAACTTCAATGTTGAT 1743
      |||
2802 erMetSerAsnSerValSerIleSerIleSerIleSerIleSerIleSerIle 2818
      |||
1744 CAAATTTATCTCTCTCAATTTGCTAGGAGATCTCAGCATTTGAGCTA 1793
      |||
2819 AlAserValSerIleSerIleSerIleSerIleSerIleSerIleSerIle 2835
      |||
1794 CGGCTGGCGCCCTGGAGCGGTTCTCAAGATGAGCTACGCGCTGCGCTC 1843
      |||
2835 hrSerMetSerIleSerIleSerIleSerIleSerIleSerIleSerIle 2851
      |||
1844 CGGGTGGCTCTA.....GCCAGCATTTGAGCTACGCGCTG 1878
      :::::::::::::::::::: |||
2852 SerThrSerMetSerAsnSerValSerIleSerIleSerIleSerIleSe 2868
      |||
1879 CGCCCTGGCAGCGGTAGCCCAAGATTGGAGCTACGCGCTGCGGCTG 1928
      |||
2868 rAlAserThrSerMetSerAsn.....SerValIle 2878
      |||
1929 ATCTGACATTTGAGCTACGCGCTGCGCGCTGCGAGCGGTTCTCAAGAT 1978
      :::::::::::::::::::: |||
2878 erAlAserIleSerIleSerIleSerIleSerIleSerIleSerIleSer 2887
      |||
1979 GGAGCTACGCGCTGCGCGGCTGCGCTGCTACGACCATTTGAGCTACGCG 2028
      :::::::::::::::::::: |||
2888 ...SerThrSerIleSerIleSerIleSerIleSerIleSerIleSerIle 2903
      |||
2029 CTGCGCCCTGGCAGCGGTAGCCCAAGATTGGAGCTACGCGCC 2068
      |||
2903 nSerIleSerIleSerIleSerIleSerIleSerIleSerIleSerIleSer 2916
      |||
seq_name: sp_bacteria:052781
seq_documentation_block:
ID 052781 PRELIMINARY; PRT; 1112 AA.
AC 052781;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-1998 (TREMBlrel. 06, Last annotation update)
DE SURFACE LAYER PROTEIN.
OS Campylobacter fetus.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBL_TaxID:196;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CIP 5396T;
RA Guesdon J.L., Casademont I., Chevrier D.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF048699; AAC04580.1; -;
SQ SEQUENCE 1112 AA; 112503 MW; DA9D9C2F62602091 CRC64;

```

alignment_scores: Quality: 190.50 Length: 666

```

Ratio: 0.540 Gaps: 33
Percent Similarity: 53.003 Percent Identity: 23.423
alignment_block:
US-09-306-689-12 x 052781
Align seg 1/1 to: 052781 from: 1 to: 1112
277 CGCCCT.....GGCAGCGGTAGCCCAAGTTGGAGCTACGCGCTGCGTCC 320
      |||
17 ArgProThrGluGlyGluGlyThrIleuAsnAsnGluSerPheVal 33
321 GGGTGAATCTACCTCCCAAAACTGGGGCAAAAATATATCTCTATA 370
      ::|
33 AsnGlyTyrPheIleGlyThrSerIleValIleuAsnIleuAsnIleuAspI 50
371 TTCGCCAAATTAATCAATATGATPACTGAAACAGTAAATGCT.....TTA 414
      |||
50 leThrLysSerPheLeuGlyIuThrIleuAsnAsnGluSerPheVal 66
415 CAGGATTTAGTCAAGCGCCGAGAGATTGGGATTTGAGTCAAGAGAGA 464
      ::|
67 LysHisIleuPheLysAsnAlaValAsnIleuThrGluPheValSerAspI 83
465 AGAGCGCAATATATTGCAACAGCTCAAAACAGTTTAGCGACGATTCMAA 514
      |||
83 uGlnLys.....AlaGlyLeuYST 90
515 CGGCTATTTGGCTTAACGTAGCGCTGCGCATTTGTTATCCGCTCCACAAT 564
      ::|
90 yTrpValAsnIleuAsnAsnGlyThrValSerLysIleAsnIleuVal 106
565 GATTAATTTGCTACAGAACTTAAGCAGCGCCAGCATTTAGTTGCGCGA 614
      ::|
107 GlyHisPheValAsnAlaValAsnAspProSerAsnAlaGlyAlaAsnG 123
615 AAGCATTTGACAAATATGCAAAATTAAGCCCAAACTGATTAATGCGATTC 664
      ::|
123 nAspIleuPheAsnAsn.....LysValIleValSer..... 133
665 AATCTATTTTGGCTCAGTATTTGGCTGGAATGATTTGAGCTACGCGCTTA 714
      ::|
134 ..AsnTyrValAlaIleAspThrIleAlaLysLeuProIleuAsnIleuThr 149
715 CAGATTAACAGCAACACATCTCTGCTAAGAGCTGCGTTGGAGCTAAC 764
      ::|
150 ProAsp.....GlnGlnAsnAlaLeuIleGlnLysThrValAspIleI 164
765 AATTCATTAATTGAATAATGCTAATTCAGTAAANA..... 804
      |||
164 eAsnAsnValThrSerAsn...SerSerValIleuSerIleAlaLysGlyG 180
805 ..CTTGACGAATTTGGTGGAGCAAAATTTAGTCAATTTGCT..... 840
      ::|
180 lValAsnIleuLysGlyIleuSerIleAsnIleuAlaGlyIleuAsnIleuS 196
841 .....TCAAAACTACAAAATATCAAAAGC..... 864
      ::|
197 AlaLeuThrThrGluAsnAspThrIleThrGlyIleGlyIleAspLe 213
865 .....TTAGGAGCTTTGGAGCAAAACCTCAAAAATATTCGGTG 901
      |||
213 uIleSerGlyValIleGlyThrAlaIleIleuSerThrIleuAsnProGly 230
902 GACTTGATTAACGTGCGCTGTTAGATGATTTCTCAGCGCTATTATTCG 951
      ::|
230 sPlyIleAsnIleuAsnIleuAsnIleuAsnIleuAsnIleuAsnIleu 246
952 GCGCAGCAACGCTGACCTTGTACTTGCAGATAAATGCTTCAACAGCTAA 1001
      ::|
247 AsnAsnIleuLysGlyLeu.....LysAsnAspGlyTyrIleLe 259
1002 AAAAGTGGGTGGGTTTGAATTTGGCAAAACCAAGTTGTGTATATTA 1051

```

```

1052 CCAAGACGGCTTCTTACATTTAGCCCACTGTTGCA..... 1092
275 hrpheaspala1ysg1ylleasplyleuglnthrvalala1leusercly 291
1093 ...GCAGGTTATCTCAACTGGCGCTGCTGCTTAAATGCTTAC 1139
292 glulysg1yleservalthrsp.....leualaasn1l 303
1140 TGTTCCTCTGCGATTAGCCCATTTGCGGTAATGCCATAAT 1189
303 evalaspvalgluleuthrasnleulys.....Alasp1ysp 316
1190 TTAATCATGCAAAAGTTAGAGT...TATGCCAAGCGTTAAAAA 1236
316 heasn.....Valasperi1etryAlasp1ysvalleuasp 328
1237 TTAGGCTATGACGAGATTAATTTATAGCAAAATATCAGCGGAGACAG 1286
329 glyserAlaspvalglinsnleulysval.....Asnglyvalgl 342
1287 GACTATTGATGATCGGTACT.....GCAATTAAATCCGCAT 1324
342 yAla1ysgl1yAlaservalAla1lethrAlaasp1ys1legluthrleua 359
1325 TGGCGGCTATGCTGCTGCTGCTGCTGCTGCGGCGATTAAACATTT 1374
359 snleuasnthrthrglyser1inserphevalserAlaaspvalAlaser 375
1375 GAAAAGTTAAACATTAATCTTGTCATCAGCAATGCAAAAAAGAGAAAGT 1424
376 l1eserval1ysgl1yasnAla1asnleuserleu1arhrg1yAla1stH 392
1425 GACCATTCAAAACTGTTCGAGAGCGCTGATTTGCTAAAGACTGCCCTA 1474
392 rThrThrleuasp.....AlaserPhegl1yAla1leuasp 406
1475 ATTATAAAGCACTAAAGATGAGAAATCGAAGAAATCATCGTCAAAAT 1524
406 laasp1euserThr...SerAlaserVal1mrser1le1ysgl1yAlasn 421
1525 GGC...GAGCGGATCACCTCAAGCAAGTT..... 1551
422 glyasnsp1ys1lethr1le1yAla1valAla1asnVala1ala1leas 438
1552 .....GATGATCTTATCGCAAAAGGTAAACGGCAAAATTA 1585
438 pgl1ygl1yAla1gl1yasnsp1leuval1le1ygl1ySer.....T 452
1586 CCCAAGATGAGCTATCAAAAGTTGTTGATACCTAGAA..... 1623
452 hrAlaaspThrleuglnProThrleuthrasn1legl1yval1Thrval 468
1624 .....TTGCTCAACATACCAAAATGT 1646
469 Asp1yasnThr1y1yasp1euthrleuserleulysAla1aglnserVa 485
1647 GACAACACGCTTACATAAGTAATCTCATCTGTAAGTCATTTACCTCGT 1686
485 lThrg1userSerPhe1yasn1leAla1y1Thrval1Thrg1userasn 502
1697 CTAT...GATTGAGAAATGATTAGT.....GCTCCA 1728
502 l1yasnvalgl1uthrvalasn1leleuAla1asnAla1arh1yasp1ysAla 518
1729 ACTTCATGTTGATCAAAAGTTATCTCTCTCAATTTGCTAGGAGATC 1778
519 Val1hr1leasnasp1y1userleulysThr1leasnPheseraspvalas 535
1779 TCACGATTGACGCTACGGCTGCGCCCTGCGAGGGGTTCTCAAGATTGGA 1828

```

```

535 pasPlys.glyAlaSerValarAla1y1sgly.....Lys1leVal 548
1829 GCTACGGCGCTGCTCGCGGCGCTCTAGCCAGCTAGGCTAGGCGCTG 1878
549 Alaasp1ysAla1thr1g1leuthr1leasnSer1n1ysval1hrAla1 565
1879 CGCCCTGCGACCGCTAGCCAGATTGAGCTACGCGCTGCGTCCGGCTG 1928
565 Ala1aspAla1valValgl1nAla1a1asnAla1arh.....Lys1leA 579
1929 ATCTC.....AGCATTTGAGCTACGCGCTCGCCCTGGCA 1963
579 sp1leasnAla1a1Ala1yasp1hrValgl1yleuthr.....Leugly 592
1964 GCGGTTCTCAAGATT...GGAGCTACGCGCTGCTCGCGGCGCTAGC 2010
593 glyvalAla1y1leuthr1yasp1euthrval1asn1n1ysgl1yAla1p1eal 609
2011 CAGCATTTGAGCTACGCGCTGCGCTGCGGCGCTAGCAGATT 2056
609 aleuthrGlyAla1asn1arThrasp1eulaspSerVal1y1asnleu 624

```

seq_name: sp_rodent:Q61869

```

seq_documentation_block:
ID Q61869 PRELIMINARY; PRT; 707 AA.
AC Q61869;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE KERATIN COMPLEX 2, BASIC, PROTEIN 2 (KERATIN 2 EPIDERMIS).
GN KRT2-17 OR KRT2-2 OR MK2E.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI;
RX MEDLINE=94149286; PubMed=7508961;
RT "The large type II 70-kDa keratin of mouse epidermis is the ortholog
RT of human keratin K2e."
RL J. Invest. Dermatol. 102:165-170(1994).
CC -1- SIMILARITY: TO ALL OTHER INTERMEDIATE FILAMENT PROTEINS.
DR EMBL: X74784; CAA52788.1; -.
DR MGI: 96699; Krt2-17.
DR INTERPRO: IPR001664; -.
DR INTERPRO: IPR003054; -.
DR PFM: PFO0038; filament; 1.
DR PRINTS: PR01276; TYPE2KERATIN.
DR PROSITE: PS00226; IF; 1.
KW Keratin; Intermediate filament; Coiled coil; Heptad repeat pattern.
SQ SEQUENCE 707 AA; 70977 MW; D58FA5E6E30BB86 CRC64;

```

alignment_scores:

Quality:	187.50	Length:	750
Ratio:	0.593	Gaps:	38
Percent Similarity:	42.133	Percent Identity:	21.867

alignment_block:

US-09-306-689-12 x Q61869 ..

Align seg 1/1 to: Q61869 from: 1 to: 707

```

91 GGTGCTCTACAGCATTTGAGCTACGCGCTGCGCCCTGCGACGGGTAG 140
|||||
86 glylyserSerPhegl1y1ySerSerGlyPhegl1ygl1yAla1yAla1yph 102
141 CCAAGATTGACGCTACGGCTGCGCTCGCGGTGATCTCAGCATTTGG... 186
102 e.....glylygl1yAla1y1yPhegl1y 110

```

```

187 .. AGCTACGGCTCGCCCTGGCAGC.....GTTCTCAAGATTGG 225
    ||| ||||| ||||| ||||| ||||| |||||
110 LysSerGlyPheGlyGlyGlySerGlyPheGlyGlyGlnGlyPhe 126
    ::||| ||| ||||| ||||| ||||| |||||
226 AGCTACGGCTCGCTCGCGGTGCTTACCAAGATTGGAGTCAGCGCT 275
    ::||| ||| ||||| ||||| ||||| |||||
127 GlyGlyGlySerArgPheGlyGlySerGlyPheGlyGlyGlyPhe 143
    : ||||| ||||| ||||| ||||| |||||
276 GCGCCCTGGCAGC.....GTAAGCCAAGATTGGAGCTACGGCTCGCTC 319
    : ||||| ||||| ||||| ||||| |||||
143 eGlyGlyGlySerPheGlyGlyArg.....PheGlyGlyGlyPhe 157
    : ||||| ||||| ||||| ||||| |||||
320 CGGGTGA.....TCTAGCTTCCCAAAACCTGGGGCAAAAAA 357
    ||||| ||||| ||||| ||||| |||||
157 roGlyGlyPheGlyGlyProGlyGlyPheProGlyGlyGlyIleHisGly 173
    ||||| ||||| ||||| ||||| |||||
358 AAT.....ATCCTATATATCCCAAAATTACCAATATGATAC 395
    ||| ||| ||| ||| ||| |||
174 ValSerValAsnGlnSerLeuGlnProLeuAspValLysValAspPr 190
    ||| ||| ||| ||| ||| |||
396 TCACACAGTAATGCTTTACAGAGATTACTCAAGCGCCGAAGAGTTGG 445
    ||| ||| ||||| ||||| |||||
190 oGluIleGlnAsn.....ValLysSerGlnGluArgGluG 202
    ||||| ||||| ||||| ||||| |||||
446 GGATTGAGGTACAAAGA..... 462
    ||||| ||||| ||||| ||||| |||||
202 IuIleLysThrLeuAsnAsnLysPheAlaSerPheIleAspLysValArg 218
    ||||| ||||| ||||| ||||| |||||
463 .....GAAGAACGCAATATATTCACACAGCTCAAAACAGTTTA..... 501
    ||||| ||||| ||||| ||||| |||||
219 PheLeuGlnGlnGlnAsnGlnValLeuArgThrLysTrpGluLeuGln 235
    ||||| ||||| ||||| ||||| |||||
502 .....GGCAGCATTCAAA 514
    ||||| ||||| ||||| ||||| |||||
235 nGlnLeuAlaCysArgGlnProHisIleLysProAspProIlePheGlnA 252
    ||||| ||||| ||||| ||||| |||||
515 CGGTAATGGCTTAACGTAGCGGTGATGTGTTATCCGCTCCACAAT 564
    ::||| ||||| ||||| ||||| |||||
252 IatyrIleGlyMetLeuLysLys.....GlnVal 261
    ||||| ||||| ||||| ||||| |||||
565 GATTAATTTG...CTACAGAAACTAAAGCAGCAGCATTTAGGTTCTGCG 611
    ||||| ||||| ||||| ||||| |||||
262 AspArgLeuSerAlaGlnArgThrSerGlnGlnSerGluLeuAsnAsnIle 278
    ||||| ||||| ||||| ||||| |||||
612 CGAAAGCATTTGACAAAT.....GCAAAATA 637
    ||||| ||||| ||||| ||||| |||||
278 tGlnAspLeuValGluAspPheLysLysLysTrpGluAspGluIleAsnL 295
    ||||| ||||| ||||| ||||| |||||
638 AAGCCAAACTGATATTATCTGCATTCATCTATTTTAGGCTCAGTAATG 687
    ||||| ||||| ||||| ||||| |||||
295 yArgThrSerAlaGlnAsnAspPheValThrIle..... 306
    ||||| ||||| ||||| ||||| |||||
688 GCTGAATGATTGATGAGCGCTTACAGAAATTAACAGCAACCAACATGCG 737
    ||||| ||||| ||||| ||||| |||||
307 ..LysLysAspValAspSerCysTrpMetAsp..... 316
    ||||| ||||| ||||| ||||| |||||
738 TCTTCCTAAAGCTGCGTGGAGCTAAACAATTCATTAAATGAAATATTG 787
    ||||| ||||| ||||| ||||| |||||
317 .....LysThrGluLeuGlnAlaArgLeuAspIleLeuAlaGlnGluVal 331
    ||||| ||||| ||||| ||||| |||||
788 GTAATTCAAGTAAACACCTTGACGAATTTGGTGAGCAATTTAGTCAATT 837
    ||||| ||||| ||||| ||||| |||||
331 AlaAspPheLeuArgThrLeu.....TyrAspAlaGluLeuSerGlnLeu 345
    ||||| ||||| ||||| ||||| |||||
838 GGTTCAAAATCTACAAAATATCAAGGCTTAGGACTTTAGAGACAACACT 887
    ::||| ::||| ::||| ::||| ::|||
346 GlnGlnAspValThrAspThrAsnValIleLeuSerMet...AspAsnIle 361
    ::||| ::||| ::||| ::||| ::|||
888 CAAAATATCGGTGACCTTGATAAAGCTGCGCTTGGTTAGATGTTATCT 937
    ::||| ::||| ::||| ::||| ::|||
361 rArgAsn.....LeuAspLeuAspSerIleI 370

```

```

938 CAGGGCTATTATCGGGCGCAACAGTGCATTTGACTGACATAAAAAT 987
    ::|
370 IeAla..... 371
988 GCTTCACAGCTAAAAAGTGGTGGGCTTTGAAATTTGGCAAAACCAAGT 1037
    ||||| |||||
372 .....GluValGlnAsnGln.. 376
1038 TGTGGTAAATATTCACAAAGCGTTTCTTTACATTTTAAGCCAAACGTC 1087
    :
376 ..... 376
1088 TTCACAGAGTTTATCTCAACTGGCGCTGTGGCTGCTTAAATGCTCTCT 1137
    :
376 ..... 376
1138 ACTGTTTCTTTCGATTAAGCCATTAGCATTTGCCGATTTCCGATTA 1187
    ::||| ||||| ||
377 .....TyrGluMetIleAlaHisLys 383
1188 ATTTAATCATGCCAAAAGTTTAGAGATTATGCCGAACGCTTAAAAAAT 1237
    : ||||| ||||| ||||| |||||
383 s...SerLysAlaGlnSerGlnGluLeuThrLysSerLysTrpGluGlnL 399
1238 TAGGCTAT.....GACGAGATTAATTTATACAGAAATATCAG 1275
    || ||||| ||||| |||||
399 eGlnValThrAlaValLysHisGlyAspSerLeu..... 410
1276 CGGGGAACAGGCAATTTGATGATGATGATGATGATGATGATGATGAT 1325
    ||| ||||| ||||| |||||
411 .....LysGluIleLysMetGluIleSerGluLeuAsnArgThrIle 424
1326 GCGCGTATTTGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1375
    : ||||| ||||| |||||
424 eGlnArgLeuGlnGlyGluIleSer..... 432
1376 AAAAAGTTAAACATATCTTTCATCAGCAATATGCAAAAAGGAAAGTG 1425
    ::||| ||||| |||||
433 .....HisValLysLysGlnCysIle 439
1426 ACCATTCAAAAGCTGTTCCGAGAGCTGATTTTGTCAAAAGTGGCTTA 1475
    ::||| ||||| ||||| |||||
440 GlyValGlnAspSerIleAlaAspAlaGlu..... 449
1476 TTATTAAGCAACTAAAGATGAGAAAATCGAAGAAATCATCGTCAAAATG 1525
    ||||| ||||| |||||
450 .....GlnHisG 452
1526 GCGAGCGGATCACTCAAGCAAGTTGATGATCTTTATCGCAAAAGTAAAC 1575
    ||||| ||||| |||||
452 LysGlnHis.....AlaIleLysAspAlaArg 460
1576 GGCAAAATTTACC.....CAAGATGAGCT 1598
    ||||| ||||| |||||
461 GlyLysLeuThrAspLeuGlnGluAlaLeuGlnGlnCysArgGlnAspIle 477
    ||||| ||||| ||||| |||||
1599 ATCAAAAAGTTGTTGATTAATGATGATGATGATGATGATGATGATGAT 1648
    ||||| ||||| ||||| |||||
477 uAlaArgLeuLeuArgAspTrpGlnGluLeuMetAsnThrLys..... 491
1649 CAACAGCTTAGAT.....AAGTAAATTCATCT 1677
    ||||| ||||| ||||| |||||
492 ..LeuSerLeuAspValGluIleAlaThrTyrArgLysLeuGlnGly 507
1678 GTAAGTCATTTACCTGCTTAATGATTTAGAGAAATATGATGATGATGAT 1727
    ::||| ::||| ||||| ||||| |||||
508 GlnGlnCysArgMetSerLysAspHisSerAspAsnValSerValSerIle 524
    ||||| ||||| ||||| ||||| |||||
1728 AACTTCATGTTGGATCAAAAT.....TTATCTGCTCTTCATTTGCTTA 1771
    ||||| ||||| ||||| ||||| |||||
524 eThrSerSerThrIleSerSerSerMetAlaSerLysThrGlyPheGlyLys 541
1772 GGGATCTCAG.....CATTTGAGCTACGCGCTGCGCCCTGCGC 1809

```



```

579 aileasplalaileuasnllysAlaSerSerAlasnthrasparglyle.... 594
941 GGCTATTATCGGGCCACAGCTGCACCTTGACTTGACAGAT.....AAA 984
595 .....AlathraslagnleuglyleAlagnlunasnlyls 606
985 AATGCTCAACAGCTAAAAAGTGGGTGGGCTTTGAAATGCAACCA 1034
607 asplalagnlilealys.....Alagnlalasngl 617
1035 AGTTGTTGTAATATTACCAAGCCGTTCTTCTTACATTATGAGCCCAAC 1084
617 uasnllysasplglylealalysasnglnlalasplilegnleuHlaspl 634
1085 GTGTTGCAGCAGCTTTATCTTCAACTGGCGCTGGCTCTTAATGCT 1134
634 ysllysilethrasnleu.....glylleuHls 643
1135 TGTACTGTTTCTTGGCATGAGCCCATGATTCGCGGTATTCGCA 1184
644 SerMetValAlaIargAlaValaIglYsnasnthrglnlyValAlaIthras 660
1185 TAAATTAATCATGCAGAAAGTTTAGAGATTATGCCGAACGCTTAA 1234
660 nlysAlasplilealalysasnglnlalasplilealasnansnlelysa 677
1235 AATTAGCTATGACGAGATTAATTATTAGCAGAAATATCAGCGGGACACA 1284
677 snile...TyrglueuAlaglnGlnasplnHlserSeraspile 692
1285 GGGACTATGATGATCGCTACTGCTACTGATTAACCGCATTCGCCCTAT 1334
693 lystsrlleu...AlalysValSerAlaAlaasnthrasparglyleAla... 707
1335 TECTGTGTGTCTGCTGCTGACCGCATTTAATGAAAAAGTTA 1384
708 .....LysasnlysalaglualasplalaserpheglnuthleuT 721
1385 AACATTAATCTTGTCAATCAGCAATAGCAAAAAGAGAAATGACCATTTCA 1434
721 hrlYasnnglnasnthrlleuileglnGlnlylualaleuValgluGln 737
1435 AACTGCTCCGAGAGCGTATTTGCTAAGAAGAGCCCAATATTATAAC 1484
738 Asn.....LysAlaIleasnglnleuGlnlypheAla 750
1485 AACTAAGATGAGAAATGCAAGAAATCATCGTCAAAATGCGAGCGCA 1534
750 hlsAlasplvalglnasplysglnleleu...GlnasnglnAlaAspl 766
1535 TCACCTCAAGCAGTGTGATCTTATGCAAAAAGTAAAGCGCAAAATT 1584
766 lethThrasnlysthr.....AlaIle 773
1585 ACCCAAGATGACCTATCAAAATGTTGATTAATCTATGCTGCTCAAA 1634
774 GluGlnAsnIleasnargrThrValAlaasnglypheglnleuIleuLys 790
1635 TAGCAAAAATGTGACAAACAGCTTAGATTAAGTAAATCATCTGTAAGTG 1684
790 nlysAlaglylealalathrasnlysglnleuIleu..... 802
1685 CATTTACTCTGTAATGATTCGAGAAATGTATTATGCTGCTCAACTTCA 1734
803 .....LeuGlnasnasparGleuasnarglyleasnglnuthrasnasn 816
1735 AATGTTGATCAAGATTATCTCTCTCAATTTGCTAGGGAATTCACAGA 1784
817 HlSGlnasplnlylsleasplnleuGlyTyraIleuLysgluGln 833
1785 TTGAGAGCTAGCGCTGCGCGAGCGGTCTCAAGATTGAGAGCTAG 1834
833 yglInHlSPheasnasnarglyleSerAlaValaIgluarGlnthralagly 850

```

```

1835 GCCTG.....CGTCCGGGTGGCTCT 1854
850 lylleAlasnalaleAlaIlealathleuProSerProSerArgAla 866
1855 AGCCAGCAT.....TGAGCTACGCGCTGCGCCCTGCGAGCGGTAG 1895
867 gLyglInHlSHlSValleuPheglnyserylYrHlSasnglylnAlaI 883
1896 CCAAGAT...TGAGCTACGCGCTGCGCGGTGAGTATGACATTTGA 1942
883 aValSerleuGlyAlaIalaglyleuSeraspHrthglylserThryL 900
1943 GCTAGCGCTGCGCGCTGCGAGCGGTCTCAAGATTGAGC.....TAC 1986
900 yslleGlyleu.....serThraspalagly 909
1987 GGCCTGCTCGCGGT...GCCTTACCGCAGCATTTG 2019
910 GlyleuSerGlyGlyValaIglylserTyraIgtTP 921

seq_name: sp_invertebrate:Q9VTR6
seq_documentation_block:
ID Q9VTR6 PRELIMINARY; PRT: 1039 AA.
AC Q9VTR6.
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-OCT-2000 (TRENBLREL. 15, Last annotation update)
DE CG5700 PROTEIN.
GN CG5700.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blazey R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle B., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Beeson K.Y., Baas A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottler P.,
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cusum S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris A.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalalet M., Kalush F., Karpen G.H., Ke Z., Kenton J.A., Ketchum
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

```

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003543; AAF4980.1; -;
 DR FLYBASE: FBgn0036231; CG5700.
 SQ SEQUENCE 1039 AA; 102635 MW; 6A710871139BD01F CRC64;

alignment_scores:
 Quality: 185.50 Length: 629
 Ratio: 0.602 Gaps: 31
 Percent Similarity: 48.967 Percent Identity: 22.258

alignment_block:
 US-09-306-689-12 x Q9VTR6 ..

Align seg 1/1 to: Q9VTR6 from: 1 to: 1039

```

37 TAGGGCTGGCCCTGGACGGGTCTCAAGATTGGAGC..... 75
323 TyrGlySerGlnProGlyIleGlyGlnThrGlyAlaGlyInProGly 339
76 .TAGGGCTGCTCCGGGTGCTTACCCAGCATTTGGAGCTAGCGCTGC 124
339 yTyrglyAlaGlnProGlyPheGlyGln...LeuGlyTyrglyAsnG 355
125 GCGCTGGACGGGTAGCCAAAGATTGGAGC.....TAGGGCTG 162
355 InProGlyValGlyGlyGlnThrGlyAlaGlyInProGlyTyrglySer 371
163 CGTCCGGGT.....GGATCTCAGCATTTGGAGCTGCGCT 197
372 GlnProGlyValGlyGlyGlnThrGlyAlaGlyInProGlyTyrglyVa 388
198 GCGCCCTGGACGGGTCTCAAGATTGG..... 225
388 IleProGlyPheGlyGlyGlnProGlyIleGlyGlnThrAlaAlaG 405
226 .....ACCTAGCGGCTGCTCCGGGT...GGCTTACCCAGCATTTGG 264
405 lYlYsProGlyTyrglyGlyGlnProGlyIleGlyGlySerPro..... 419
265 AGCTAGCGGCTGCGCCCTGGACGGGTAGCCAAAGATTGG..... 303
420 ValTyrglyThrGlnGlnGlyThrGlyGlyInSerGlyIleSerGly 436
304 .....AGCTAGCGGCTGCTCCGGGTGATCT..... 330
436 yGlnProGlyTyrglyThrGlnProGlyGlnThrGlyAlaGlyInProG 453
331 .....ACCTTCCCAAAAACCTGGGCAAAAAAATTATC..... 363
453 lYTyrglySerLeuProGlyThrGlyGlnAlaThrAlaGlyInPro 469
364 CTCATATTTCCCAAAATATACAA..... 387
470 GLyTyrglyProGlySerGlnProGlyIleGlyGlyGlnThrValGly 486
388 .....TATG 391
486 yHisGlyTyrglySerGlnProGlyIleGlyGlyAlaProValTyrg 503
392 ATACTGACAAGTAATGTTTACAGATTATTCAAAGCGCGAA... 438
503 lYThrGlnProGlyGlyGlyGlnThrGlyAlaIleGlyGlyGlnPro 519
439 ...GAGTTGGGATTGAGTACAAGAAAGCAATATATATTGCAAC 485
520 GLyGlnIleGly.....AspArgValGlyInProGlyTyrglyTh 533
486 AGCTCAAAACGATTTAGCAGC.....ATTCAAA 514

```

```

533 rGlnThrGlyGlnIleGlyAlaProGlyArgTyrrhrAspGlySerGlnP 550
515 CCCTATTGGCTTAACGACCGTGCATTTGTTATCCGCT.....CCA 558
550 hrValProGlyAlaValGlyThrGlyGlyValAlaAlaGlyThrSer 566
559 CAATTTGATTAATTTGCTTACGAAACTAAAGCA.....GCCCA 596
567 GLyAlaAspAspAlaPheSerGlnAlaGlySerSerIleGlyAspGly 583
597 ACCTATTAGTTTCCGCAAGCATTTGCAAAATGCAAAATTAAGCCAAA 646
583 nAlaSerAlaSerAlaGlnGly...LysLysAsnGlyGlyThrAlaLys 599
647 CTGATTTATCGCATTCATTCATTTTATAGCTCAGATTGCTGCGATG 696
599 hrGlnValSerGlyThrTyrrSerSerGlyGlyThrPheSerAla..... 613
697 GATTTAGATGAGCCTTACAGAAATACAGCAACCAACATGCTTGTCTAA 746
614 .....SerAlaMetThrSerAspAlaAspArgAlaAlaSerAlaG 627
747 AGCTGCTTGAGCTTACCAAAATTCATTAAATGAAATATTGCTAATTCAG 796
627 nValThrGlyAsnAlaAspGlyAlaValSerGlnSerGlnGlySerGly 644
797 TAAAAACCTTGACGAAATTTGGTGACCAATTTGTCATTTGCTTCAAAA 846
644 lYProAlaGlnSerGlnAlaGlnValAlaGlnAlaLysAspGlyGlyThr 660
847 CTACAAAATATCAAAAGCTTAGGCACTTTAGGACAACTCAAAATAT 896
661 LysAlaSerSerGlnSerGlyIleIleGlnGlnSerGlnSerGly 677
897 CGGTGACTTGATTAACCTGCGCTTGTAGATGTTATCTCAGGCTAT 946
677 lHisAlaAsnAspLysGlyGly..... 684
947 TATGCGGCGCAACGCTGCACTTGTACTTGCAGTAAATATGCTTCA... 993
685 .....LeuAlaAspAlaGlnSerSerGly 692
994 .....ACAGCTAAAAAAGTGGTGGCGGTTT.....GATT 1025
693 ProGlyGlnThrSerSerGlnAlaGlnIleGlyPheArgProGlyGln 709
1026 GGCAAACCAAGTGTGTGTAATATTACCAAGCCGTTCTTCTTACATTT 1075
709 uAlaAsnProIleAlaAlaAsnGlyGlyGlnAlaSerSer..... 723
1076 TAGCCCAACGTTGTGACAGGTTTATCTTCACTGAGCGCTGCTGCT 1125
724 .....SerSerGlyThrHisSerSer...GlnSerSerSer 734
1126 TTAATGCTTCTACTGTTCTCTTGGATTAGCCCATTTAGCATTTGCCG 1175
735 GlnIleHisGlyThrSerSerPheGlyValSer.....TyrHisG 748
1176 TATTGCCGATTAATTTAATCATGCAAAA...AGTTTAGAGAGTTATCCG 1222
748 yAlaAlaGlnSerAlaSerGlyThrLysGlnGlnValAlaThrTyrrg 765
1223 AAGCTTTAAAAAATTTAGGCTATGACGAGATATTTATAGCAGATAT 1272
765 lValAlaAsnArgGlnLeu.....PheAsnThrIleSerGlnPhe 777
1273 CAGCGGGGAACAGGACTATTGATGATCGATTGATTCGAATTAATACCG 1322
778 GLyAsnAsnAlaAsnAlaValAlThrAspArgAlaAspAlaValTyrr 794
1323 ATTGCGCGCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1372

```

```

794 yProalaLeuThrAspGluSerAspArgValProGluAlaGlnLeuLys 811
1373 TTGAAAGTAAACAAATCTGTCATCGACGATGCAAA.....1413
      ||| |||.....::: |||||
811 erThyrProLysGluGluAlaGlnValAlaSerLysLeuAspHis 827
1414 .....AAAGAAAGTGACCATTCCAAACTGGTT 1442
      ||| .....:::
828 ProGluProValGlnTyrAspAspGluAlaAspProAspGluTyr 844
1443 CCGAGAGCGCTGATTTTGTCTAAGAGTCCTTAATTAAGCACTAAG 1492
      : ||| .....::: |||
844 rAspGluAspGluTyrTyrGlnGluLysPro.Val.....Lys 856
1493 AAGAGAAATTCGAGAAATCATCGCTCAAAATGCG.....AGCGATC 1536
857 LeuGluGlnSerProLysSerSerThrIleValAlaProThrGluAlaThr 873
1537 ACCTCAAGCAAGTGTGATCTTATCG.....CAAA 1568
      : |||.....|||
873 rProArgProLysThrTyrAsnGlnSerSerProThrGlnSerGlnGln 890
1569 AGGTAAAGCAAAATTACCAAGATGACTATCAA 1603
      |||.....::: |||
890 hrValValAlaProValProGluGluLysTyrGln 901
seq_name: sp_vertibrate:013083

```

```

seq_documentation_block:
ID 013083 PRELIMINARY; PRT; 722 AA.
AC 013083;
DT 01-JUL-1997 (TREMBLrel. 04. Created)
DT 01-JUL-1997 (TREMBLrel. 04. Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15. Last annotation update)
DE AFEP POLYPROTEIN PRECURSOR.
OS Dissostichus mawsoni.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphi; Acanthopterygii; Percomorpha; Perciformes;
OC Notothenioidae; Nototheniidae; Dissostichus.
OX NCBI_TaxID=36200;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97268652; PubMed=9108060;
RA Chen L., Devries A.L., Cheng C.H.;
RT "Evolution of antitrypsin glycoprotein gene from a trypsinogen gene in
RT Antarctic notothenioid fish."
RL Proc. Natl. Acad. Sci. U.S.A. 94:3811-3816(1997).
DR EMBL; U58944; AAB57731.1; -.
KW Signal; Polyprotein.
FT PROPEP ? ?
FT SIGNAL 1 14 POTENTIAL.
FT CHAIN ? 722 ANTIFREEZE GLYCOPROTEIN POLYPEPTIDE POLYPROTEIN.
SQ SEQUENCE 722 AA; 65378 MW; F1F3443E0A35E018 CRC64;

```

```

alignment_scores:
  Quality: 173.00      Length: 728
  Ratio: 0.564         Gaps: 25
  Percent Similarity: 42.170   Percent Identity: 22.527

```

```

alignment_block:
US-09-306-689-12/rev x 013083 ..

```

```

Align seg 1/1 to: 013083 from: 1 to: 722
2082 TCCACCCGAGCGAGCGCTAGCTCCATCTTG.....GCTACCG 2042
      : ||| : |||.....||| |||
75 AlaThrProAlaThrAlaAlaThrProAlaLeuIlePheAlaAlaThrAl 91
2041 TCGCAGGAGCGAGCGCTAGCTCCATGCTGCTAGAGCCAGCCGAGCG 1992
      ||| : ||| : |||.....|||
91 AlaThrProAlaThrProAlaLeuIlePheAlaAlaThrAlaAlaThrP 108

```

```

1991 AGGCGTAGCTCCATCTTGAGAACCGCTGCC.....AGGCGCAG 1951
      |||.....||| |||||
108 roAlaThrAlaAlaThrProAlaThrAlaAlaPheAsnPheAlaAlaThr 124
1950 GCCGTAGCTCCATGCTGAGATCCACCGGAGCGAGCGCTAGCTCCAT 1901
      |||.....||| .....:::
125 AlaAlaThrProAlaThrAlaAlaThrAlaAlaLeuIlePheAlaAlaThr 141
1900 CTGGCTACCGCTGCCAGGCGCAGCGCTAGCTCCATGCTGCTAGAG 1851
      |||.....||| .....:::
141 rAlaAlaThrAlaAlaThrProAlaThrProAlaLeuAsnPheAlaAlaThr 158
1850 CCACCCGAGCGAGCGCGCTAGCTCCATCTTGAGAACCGCTGCCAGCG 1801
      : ||| .....||| |||||
158 hrAlaAlaThrAlaAlaThrProAlaThrProAlaThrAlaAlaLeuAsn 174
1800 CAGCGCTAGCTCCATGCTGAGATCCCTAGCAAAATTCGAAGATTA 1751
      |||.....||| .....:::
175 PheAlaThrAlaAlaThrAlaAlaThrProAla..... 185
1750 AACTTGATCCAACTGAATGAGCCACTAATACATTTCTGCAATCA 1701
      : |||.....|||
186 .....ThrProAlaLeuAsnPheAlaAlaThrAla 195
1700 TTAGAGAGTAATGCACCTTACAGATGAGATTAATCTATCTAGCTGT 1651
      : |||.....|||
196 AlaThrProAlaThrAlaAlaThrProAlaLeuAsnPheAla..... 209
1650 TGTCACTTTTCTGATGTTTGAGCAATTCATGTTATCAACACTTTTG 1601
      : |||.....|||
210 .....AlaThrAlaAlaThr 214
1600 ATAGCTCATCTGGTAATTTGCCCTTACCTTTGCGATTAAGATCA 1551
      : |||.....|||
214 hrAlaAlaThrProAlaThrAlaAlaLeuAsnPheAlaAlaThrAlaAla 230
1550 ACTTGCTTGAGTGATCCGCTGCCATTTTGACCGATGATTTCTTCGAT 1501
      : |||.....|||
231 ThrAla..... 232
1500 TTTCATCATCTTAGTGTCTTATTAATTAAGCACTCTTACCAAAATCA 1451
      : |||.....|||
233 .....AlaThrProAla 236
1450 CCTGTGGAACAGTTTGAATGTCATCTTCTTTTGTATTCGTG 1401
      : |||.....|||
236 lathr.....AlaAlaPheAsnPhe..... 242
1400 ATGACAAGATTAGTTAACTTTTCAAAATGTTAAATCGGCTGCAGCAG 1351
      : |||.....|||
243 .....AlaAlaThrAla 246
1350 AGACACACCAACCAATAGCGGCAATGCGGATTAATTCAGTAACCG 1301
      : |||.....|||
246 AlaAlaThrProAlaThrAlaAlaThrProAlaLeuIlePheAlaAlaThr 263
1300 ATGCATCAATAGTCCTGTCGCCGCTGATATTCTGTATAATAATATCT 1251
      : |||.....|||
263 lathr.....ProAlaThrAlaAlaThr 271
1250 CGGTCAATAGCTTAATTTTAAAGCTTGCGCATTAATCTCTTAACCTTT 1201
      : |||.....|||
272 ProAla.....LeuAsnPheAlaThrAlaAlaAlaAlaAlaThrProAlaThr 287
1200 TGCATGATTAATTTATCGCAATACCGCAATGCTAATAGGCTTAATCG 1151
      : |||.....|||
287 rProAlaLeuAsnPheAlaAlaThrAlaAlaThrAla.....AlaThrAlaAla 303
1150 CAAGAGAAACAGTAGAAGCAATTAAGACCAAGCCAGCCGATTTGAAGAT 1101
      : |||.....|||
303 lathrProAlaThrAlaAlaLeuAsnPheAlaAlaThrAlaAlaAla 319

```

```

1100 AACCTGCTCAACACGTTGGGCTAAATGTAGAGAAAGCGCTTGGT 1051
    |||||
320 ThrProAlaThr..... 323
1050 AATATTACCAACACTGGTTGGCAATTCAAAACCGCACCTTTT 1001
    |||||
324 .....ProAlaLeuAsnPhaAlaAlaThrProAlaThr 338
1000 TAGCTGTGAAGCTTTTATCTGCAGTCAAGTGA..... 963
    |||||
338 AlaAlaThrProAlaLeuMetPhaAlaAlaThrAlaAlaThr 354
962 ...GCTGTGGCCCGCATATAGCCCTGAGATACATCTAAACCAAGGCC 916
    |||||
355 ProAlaThrAlaAlaPhaAsnPhaAlaAlaThrProAlaThr 371
915 AGCTTTTCAGAGTCCACCGATATT.....TTGAGTTTGCTC 878
    |||||
371 rAlaAlaThrProAlaLeuIlePhaAlaAlaThrAlaAlaThrAla 388
877 CTAAAGTCCCAAGCCTTGATATTGTGAGTTGAACCA..... 837
    |||||
388 hrAlaAlaThrAlaAlaLeuIlePhaAlaAlaThrAlaAlaThr 404
836 .....AATTGACTAATTGCTACCAAAATTCGTGAC 805
    |||||
405 ProAlaThrAlaAlaPhaAsnPhaAlaAlaThrAlaAlaThrProAla 421
804 TGTATTACGAAATTAGCAATATTTCATTAATGAATTTGTGCTCCA 755
    |||||
421 rAlaAlaThrProAlaLeuIlePhaAlaAlaThrAlaAlaThrAla 438
754 AGCAGCTTTGCAAGACATGTTGTTGTTCTGTAATCTGTGAAGCCCTCA 705
    |||||
438 hrProAlaThrAla.....AlaLeuAsnPhaAlaThrAlaAla 450
704 TCTAATTCATTCAGCCCAATCTAGCCCTAAATAGATTGAATGCCAGA 655
    |||||
451 ThrAlaAlaThrProAlaThr..... 457
654 TAATACAGTTTGGCTTATTGTCATTTTGACAAATGCTTCGCAGAAC 605
    |||||
458 .....AlaAlaLeuAsnPhaAlaThrAlaAlaThrAlaAla 470
604 CTATGCTGGCGCTTACTTGTGCTAGCAATTAATCAATTTGTGGA 555
    |||||
470 hrProAlaThrProAlaLeuAsnPhaAlaAlaThrAlaAla..... 483
554 GCGGATAACACAAATGCCAGCTCAGTTAAGCAATAGCGTTGAATCGT 505
    |||||
484 .....Th 484
504 GCGTAACCTGTTGAGCTGTGCAATTAATTCGCTTCTCTTGTGA 455
    |||||
484 rProAlaThrAlaAlaThrProAlaLeuMetPhaAlaAlaThrAlaAla 501
454 CCGCAATGCCCAACCTTCGGCGCTTGACTAATCCCTGAACCAATTA 405
    |||||
501 hrAlaAlaAlaThrProAlaThrAlaAlaPhaAsnPhaAlaAla..Thr... 515
404 CCTGTTCAGTATCATATTGTAATTTGGGAATATAGAGATAATTTT 355
    |||||
515 ..... 515
354 TTTTGGCCCGACTTTTGGGAAGTAGATCCACCGGACGACGCCGTAGC 305
    |||||
516 .....AlaAlaThrProAlaThrAlaAlaAlaThr 524
304 TCCCAATCTG.....GCTACGCTGCAGG...GCGAGCGCCGTAG 267
    |||||
524 rProAlaLeuIlePhaAlaAlaThrAlaAlaThrProAlaThrAlaAla 541
266 CTCCAATGCTGGCTAGAGCCACCGGACGACGCCGTAGCTCAATCTTG 217

```

```

541 hrProAlaLeuIlePhaAlaAlaThrAlaAlaThrProAlaPhe 557
    |||||
216 AGAACCGCTGCCAGG.....GCCAGCGCTTACTGCTCAATGCT 179
    |||||
558 AsnPhaAlaAlaThrAlaAlaThrProAlaThrAlaAlaThrProAla 574
178 GAGATCCACCGGACGACGCGGTAGCTCCAAATCTTGCTACCGCTGCC 130
    |||||
574 uIlePhaAlaAlaThrAlaAlaThrAlaAlaThrProAlaThrAlaAla 591
129 .....AGGCGCAGGCGCTTACCTCCATGCTGCT..... 100
591 heAsnPhaAlaThrAlaAlaThrAlaAlaThrProAlaThrAlaAlaPhe 607
99 .....AGAGCCACCGGACGACGCGCTTACTGCTCAATGCT 65
    |||||
608 AsnPhaAlaAlaThrAlaAlaThrProAlaThrAlaAlaThrProAla 624
64 GAGAACCGCTGCCAGGCGGACGCGGTAGCT 34
    |||||
624 uMetPhaAlaAlaThrAlaAlaThrProAla 634

```

seq_name: sp_bacteria:Q9SOT3

```

seq_documentation_block:
ID Q9SOT3 PRELIMINARY: PRT: 565 AA.
AC Q9SOT3:
DT 01-MAY-2000 (TRENBLREL, 13, Created)
DT 01-MAY-2000 (TRENBLREL, 13, Last sequence update)
DT 01-JUN-2000 (TRENBLREL, 14, Last annotation update)
DE FLAGELLIN.
GN FLIC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=U1A-44;
RA Ohnishi K., Ishioke K., Matsuda T., Harayama S.:
RT "Cloning of H antigen genes in E.coli serotypes and expression in
RT E.coli K-12."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB028480; BAA85089.1; -
DR INTERPRO: IPR001029; -
DR INTERPRO: IPR001492; -
DR PFAM: PF00669; Flagellin_N; 1.
DR PFAM: PF00700; Flagellin_C; 1.
DR PRINTS: PR00207; FLAGELLIN.
SQ SEQUENCE 565 AA; 58528 MW; 68EFB39B66436342 CRC64;

```

```

alignment_scores:
  Quality: 170.00 Length: 543
  Ratio: 0.639 Gaps: 25
Percent Similarity: 48.987 Percent Identity: 22.284

```

alignment_block:

US-09-306-689-12 x Q9SOT3 ..

Align seg 1/1 to: Q9SOT3 from: 1 to: 565

```

469 CGCAATATATTGCAACAGCTCAACAGCTTACGACAGATCAACCGC 518
    |||||
15 GlnAsnAsnIleAsnLysAsnGlnSerAlaLeu.....SerSerSe 28
519 TATTGGCTTAACTGAGCGGTGATGTTATTCGCTCCACAAATTTGATA 568
    |||||
28 rIleGlnArgLeuSerSerGlyLeuArgIleAsnSerAlaLysAspSp. 44
569 AATTGCTACAGAAAACCTAAAGCAGGCCAAGCATTA..... 603
    |||||

```


[illegible]

```

1204 AGTTAGGAGAGTATTGGCCGACCGCTTTAAAATAATAGCATATGCAGACA 1253
    ||| : : : : : |||
351 SerThrAsnGlyGlyValLysPheIysAspThr...ValSerSerAs 366
1254 TAATTATTATTCAGAAATATCAGCGGGAGACAGGACTATTGATGATCCGG 1303
    | ||||| : : : : : 
366 palLeuAlaInglValLysAlaAspSerThrAlaAsnValLysI 383
1304 TTACGCAATTAAATCCGCAATG.....GCCGATATGCTGGTGT 1344
    : ||| ||| ||| : : : : : 
383 IeThrPheAsnAsnGlyProLeuSerPheThrAlaSerPheGlnAsnGly 399
1345 GGTGTGCTGTCGCGCCGATTAATCATTTTGAAAAAAGTTAAATCAATGCT 1394
    ||||| : : : : : |||
400 ValSerGlySerAlaAlaIser.....AsnAl 408
1395 TGTGCATCACGATATAGCAAAAAAGAGAAGATGACCATTCAAACTGGTTCC 1444
    : : : : : ||| : : : : : 
408 aLaIleTryIleAspSerGluGlyGluLeuThrTrfHglUserTy.... 423
1445 GAGAGCGCTGATTTTGCTTAAGAACATGCGCTTAATTTAAAGCACTAAAGAT 1494
    : ||||| : : : : : |||||
424 .....AsnThrAsnTySerValAspLysAsp 432
1495 GAGAAATATCGAAGAAATCATGCTGCTCAAAATGGCGAGGATGACCCCAAA 1544
    : : : : : ||| : : : : : 
433 ThrGlyAlaValSerValThrlGlyGlySerLy.....ThrIlyLy 446
1545 GCAGATGATGATCTTATTCGCAAAA.....GGTAAAGCGCAAA 1582
    | : : : : : ||| : : : : : 
446 sTryrAlaAlaAsnValGlyAlaGlnAlaTyGValGlyAlaAspGlyLyl 463
1583 TTACCCAAGATGAGCTATCAAAAGTTGTTGATACATGCAATTGCTCAAA 1632
    : ||| : : : |||
463 eUrThrAsnThrSerThrlLy.....Ser 472
1633 CATGACCAAAATGTGACAACACAGCTTGATAGTTATGTCATGCTAGTAG 1682
    : : : : : ||| : : : : : ||| : : : : : 
473 AlaThrLysAspProLeuAsnAlaLeuAspGlnAlaIleAlaSerIleAs 489
1683 TGCATTTTACSTGCTGTATGATTCGAGAAAATGATTAGTGCGCTCCAACTT 1732
    : ||| |||||
489 pLysPheArgSerSer.....LeuGlyAlaIleGlnA 500
1733 CAATGTGATCAAAAGTTTATCTCTCTT 1761
    : : ||||| : : : : : |||
500 snArgLeuAspSerAlaValThrAsnLeu 509

seq_name: sp_bacteria:Q9RY75

seq_documentation_block:
ID Q9RY75 PRELIMINARY; PRT: 1467 AA.
AC Q9RY75;
DT 01-MAY-2000 (TREMBLE). 13, Created)
DE 01-MAY-2000 (TREMBLE). 13, last sequence update)
DE 01-JUN-2000 (TREMBLE). 14, last annotation update)
DE CONSERVED HYPOTHETICAL PROTEIN.
DR DRO075.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Elsen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
  Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
  Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
  Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
  Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
  Ketchum C.M., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
  Fraser C.M.;

```

RT "Genome Sequence of the Radioreistant Bacterium Deinococcus
 RT radiodurans R1.";
 RL Science 286:1571-1577(1999).
 DR EMBL: AE001870; AAF09665.1;
 DR TIGR: DR0075;
 DR INTERPRO: IPR00122;
 SQ SEQUENCE 1467 AA; 148507 MW; 8964BA26A82693FF CRC64;

alignment_scores:

Quality: 169.00 Length: 838
 Ratio: 0.426 Gaps: 40
 Percent Similarity: 47.375 Percent Identity: 22.196

alignment_block:

US-09-306-689-12 x Q9RY75 ..

Align seg 1/1 to: Q9RY75 from: 1 to: 1467

```

43 CTGGCCCTGGCAGCGCTTCT...CAAGATTGAGCTACGGCTGCGTCC 89
||||| ||||| : : : ||||| : : : |||||
194 LeuArgAspGlyAlaThrAspLeuGlnGlyThrValLysLeuArgAs 210
90 GGGTGGCTCT.....AGCCAGCATTTGAGCT 115
||||| : : : : : : : : : : : : :
210 pGlyAlaSerAspLeuAlaAspGlyAlaAsnArgAlaAlaAspGlySerA 227
116 ACGGCTGCGCCCTGGCAGCGTAGCCAAAGATTGAGCTACGGCTGCGT 165
||||| ||||| : : : ||||| : : : |||||
227 rGlySerAlaAspGlyAlaGlyThrLeuSerSerGlyValGlyArgLeu 243
166 CCGGGGTGATCTCAGCATTTGAGCTACGGCTGCGCCCTGGCAGCGATT 215
||||| ||||| : : : ||||| : : : |||||
244 ThrAspGlyThrValLysLeuSerAspGlyLeuArgGlnLeuGluAlaAl 260
216 TCAAGATTGGAGCTACGGCTGCGTCCG.....GGTGGCTTACCGCAGC 259
||||| : : : ||||| : : : ||||| : : : |||||
260 AlaProGlyGluThrGlnLeuAlaProLeuArgSerGlyAlaGlyGln 276
260 ATTGAGACTACGGCTGCGCCCTGGCAGC..... 288
||||| ||||| : : : ||||| : : : |||||
277 .....LeuArgThrGlySerGlnSerLeuAlaSerGlyLeu 288
289 .....GGTAGCCAAAGATTGAGCTACGGCTGCGTCCG.. 321
||||| : : : ||||| : : : ||||| : : : |||||
289 GlyGlnLeuAlaAspGlyThrAspGlnLeuAlaAlaGlyAlaLysAl 305
322 ...GGTGGATTCTAGCTTCCCAAAACTGGGCAAAAAAATTATCTCT 367
||||| : : : ||||| : : : ||||| : : : |||||
305 AsnGlyGlyAlaGlnGlnValAlaAlaGlyAlaAsnLysGlnLeuAlaAlaG 322
368 ATATTCCCAAAATTACCAATATGATTACTGAACAAGTAATGTTAGCT 417
||||| : : : ||||| : : : ||||| : : : |||||
322 InLeuProGln.....LeuGlnSerGlyLeuThr 331
418 GATTAGTCAAAAGCGCCGAGAGTGGGAGTTAGAGTCAAAAGAGA 467
||||| : : : ||||| : : : ||||| : : : |||||
332 GlnLeuSerSerGlyAlaAspLysLeuAla..... 341
468 ACGCAATATATTGCAACAGCTCAACAGCTTGGCAGCATTCAAACG 517
||||| : : : ||||| : : : ||||| : : : |||||
342 .....AlaGlyAlaLysSerAlaAsnGlyGlyAlaGlnGln 354
518 CTATTGGC.....TTAAGTGGCT..... 537
||||| : : : ||||| : : : ||||| : : : |||||
354 AlaAlaGlyGlyAlaAsnLysGlnLeuAlaGlnGlnLeuProLysLeuGlnGly 370
538 GGCATTGTGTATCCGCTCCAAATGATTAATTGCTACAGAAAACATA 587
||||| : : : ||||| : : : ||||| : : : |||||
371 GlyLeuThrGlnLeuSerGlnGlyAlaAspLysLeuAlaAlaGlyAlaLys 387
588 AGCAGGCCAAGCATTTAGTCTGCGGAAGCATTTGACAAAATGCAATA 637
||||| : : : ||||| : : : ||||| : : : |||||

```

```

387 sSer.....AlaAsnGlyGlyAlaGlnGlnValAlaGlyLysAlaLysG 402
638 AAGCCAAATGTTATTTATCTGCATTCATCTATTTAGCGCTGATATG 687
: : : ||||| : : : ||||| : : : ||||| : : : |||||
402 InLeuAlaAspGlnLeuProGlnLeuGlnSerGlyLeuThrGlnLeuSer 418
688 GCTGAATGATTTAGATGAGCCTTACGAATATACAGCAACAAATGCT 737
||||| : : : ||||| : : : ||||| : : : |||||
419 GlnGlyAlaAspLysLeuAlaAlaGlyAlaLysSerAlaAsnGlyAla 435
738 T.....CTGCTAAAGCTGGCTGGAGCTACAAATTCATTA..... 774
||||| : : : ||||| : : : ||||| : : : |||||
435 aGlnGlnValAlaAlaGlyAlaAsnLysGlnLeuAlaAspGlnLeuProGln 452
775 .....ATTGAATATATTGCTAAT 792
452 euGlnSerGlyLeuGlyGlnLeuSerGlnGlyAlaAspArgLeuAlaSer 468
793 TCAGTAAATGACCTTGACGAATTTGGTACGCAATTAAGT..... 831
||||| : : : ||||| : : : ||||| : : : |||||
469 GlyAlaLysAspAlaAsnAlaGlyAlaGlnGlnValAlaAlaGlyAlaAsn 485
832 CAATTGTGTTCAAACTACAAATATATCAA...GGCTTAGGACTTTAG 877
||||| : : : ||||| : : : ||||| : : : |||||
485 sGlnLeuAlaAlaGlnLeuProGlnLeuGlnSerGlyLeuGlyGlnLeu 502
878 GAGCAAAATCTCAAAATATATC...GGTGGACTTGATTAAGCTGGCTGGT 924
||||| : : : ||||| : : : ||||| : : : |||||
502 eArgGlyAlaAspArgLeuAlaSerGlyAlaLysGlyAlaAsnAlaGly 518
925 TTAGATGTTATCTCAGG..... 942
519 AlaGlnGlnValAlaAlaGlyAlaAsnLysGlnLeuAlaThrGlnLeuProGln 535
943 CTATTATCGGGCGCAGCAGCTGCATCTGTACGATTAATATGCTT 991
||||| : : : ||||| : : : ||||| : : : |||||
535 nLeuGlnSerGlyLeuGlyGlnLeuSerGlnGlyAlaAspArgLeuAla 552
992 CAACAGCTAAATAAGTGCGTGGCTGTT..... 1020
||||| : : : ||||| : : : ||||| : : : |||||
552 eArgGlyAlaLysAspAlaAsnAlaGlyAlaGlnGlnValAlaAlaGlyAsn 568
1021 ...GAATGGCAAAACCAAGTT.....GTTGGTAATAT 1049
||||| : : : ||||| : : : ||||| : : : |||||
569 LysGlnLeuAlaThrGlnLeuProGlnLeuGlnSerGlyLeuGlnLe 585
1050 TACCAAAAGCTTCTTCTTACATTATAGCCCAACGCTTGCACAGCTT 1099
||||| : : : ||||| : : : ||||| : : : |||||
585 uSerGlnGly.....AlaAspArgLeuAlaSerGlyA 596
1100 TATCTTCAACTGGGCTGTGCTGCTTTAATTGCTTACTGTTCTCTT 1149
||||| : : : ||||| : : : ||||| : : : |||||
596 IaAsnAspAlaAsnSerGlyAlaLysGlnLeuAla..... 607
1150 GCGATTAGCCCATTTAGCATTTGCCGT.....ATTGCCGATAAATT 1190
||||| : : : ||||| : : : ||||| : : : |||||
608 .....AlaGlyAlaAsnLysGlnLeuAlaAspGlnLe 617
1191 TATATGCAAAAGT.....TTAGAGAGTTATGCCAAGCT 1228
||||| : : : ||||| : : : ||||| : : : |||||
617 uProGlnLeuGlnSerGlyLeuGlyGlnLeuSerGlnGlyAlaAspArgL 634
1229 TTAATAAATTTAGCTATGAC..... 1248
: : : : : : : : : : : : : : : : : : :
634 euAlaSerGlyAlaAsnAspAlaAsnAlaGlyAlaLysGlnLeuAlaAla 650
1249 GGAGATATTTATTAGCA.....GAATATCAGCGCGGAGACAGG 1286
||||| : : : ||||| : : : ||||| : : : |||||
651 GlyAsnArgGlnLeuAlaThrGlnLeuProGlnLeuGlnSerGlyLeuGln 667
1287 GACTATTGATCATCGCTTACTGCAATTAATCCGATGGCCGCTATTC 1336
||||| : : : ||||| : : : ||||| : : : |||||
667 nGlnAlaSerAlaGlyAlaAspArgLeuAsnThrGlyAlaAlaGlnLeuAla 684

```

```

1337 CTGGTGTGCTGCTGCTGAGCCGATTATTAATTGAAAAAGTTAA 1386
      |||:::
684 laThrclythrleuInleuAsn.....AspalaVallys 655
1387 CATTAATCTTGATCATCAGATACGAAAAAGAGAAAGTGCATTCGAAA 1436
      |||:::
696 AsnGlnAlaIleuLeuProGlyInleuLysGlnGlyValAlaIalysVal 712
1437 CTGGTCCAGAGAGGCTGATTTCTTAAGACGTGCTAATTATTAAGCAA 1486
      |||:::
712 n.....SerGlyAlaMetGluValHisSerGlyThrAspT 724
1487 CTAAAGATGAGAAATCGAAGAAATCGCAAAATGGCGCGGAGTC 1536
      |||:::
724 hleuSerAsnAlaLeuProGlnleuGlnAlaGlyAlaGlyAlaVal 740
1537 ACCTCA.....AAGCAAGTTGATGATCTTATCGCAAAAGGTAA 1574
      |||:::
741 ThnGlyAlaSerGlnleuAlaGlnGlyAlaAspLysleuAlaIalGlyTh 757
1575 CGGCAAAATTTACCCAGATGAGCTATCAAAAGTTGTGATACATGAAT 1624
      |||:::
757 rAlaGlnleu..... 760
1625 TGCTCAACATAGCAAAATGCAACAGCCTTAGATAAGTTA..... 1668
      |||:::
761 .....SerGlnGlyAlaThrThrleuSerSerLysleuLysGlu 773
1669 .....ATCTCATCTGTAGTGCATTTACCTCGCTCATGTA 1703
      |||:::
774 AlaGlnGlnGlySerThrAlaAlaValGlnGlyValGlnGlnleuSerG 790
1704 TTCGAGAAATGATTTATGTCGCTCAACTTCATGTTGATCAAA..... 1746
      |||:::
790 uGlyAlaThrArgleuAlaIalGlyThrAlaGlnleuSerGlnGlyAla 807
1747 ..AGTTATCTCT..... 1758
      |||:::
807 alThrleuSerSerLysleuLysGlnAlaGlnGlnGlyThrAlaAla 823
1759 CTTCATTTGCTAGGAGATCTCAGCATTCGAGCTAGCGCTCGCC..... 1804
      |||:::
824 ValGlnGlyAlaGln.GlnleuGlyGlnGlyAlaAspLysleuAlaIal 840
1805 .....CTGCGAGCGGCTTCAAG 1822
      |||:::
840 lyThrAlaGlnleuSerGlnGlyAlaAsnThrleuSerSerLysLys 856
1823 .....ATTGAGCTAGCGGCTCGCGCTCGCTCTAGCCAGCAT 1863
      |||:::
857 GluAlaGlnGlnGlyThrThrAlaAlaValGlnGlyAla...GlnGlnle 872
1864 TGGAGCTAGCGGCTCGCGCTCGCCAGC..... 1891
      |||:::
872 uGlyGlnGlyAlaAspLysleuAlaIalGlyThrAlaGlnleuSerGln 889
1892 .....GTAGCCAAAGATTGAGCTAGC 1912
      |||:::
889 lyAlaThrThrleuSerSerLysleuLysGlnAlaGlnGlnGlyThrThr 905
1913 GGCTGCTGCGGCTGATCTCAGCATTCGAGCTAGCGCTCGCCCTGCGC 1962
      |||:::
906 AlaAlaValaIalGlnAlaGlnleuGlnGlyGlnGlyAlaAspLysleuAl 922
1963 A.....GGGTTCTCAAGATTGAGCTAGCGCTCGCTCGCGGTG 2003
      |||:::
922 baIaGlyThrAlaGlnleuSerGlnGlyAlaThrThr...leuSerThrL 938
2004 CTCTAGCCAGCATTTGA...GCTAGCGGCTCGCGCTCGCCAGCGGTAGCC 2050
      |||:::
938 yslleuLysGlnAlaGlnGlnGlySerAlaAlaIalAlaValGlnGlyValGln 954

```

```

2051 AAGATTGAGCCT 2062
      |||:::
955 LysleuAlaIalA 958

```

```

seq_name: sp_bacteria:Q9PC04

```

```

seq_documentation_block:

```

```

ID Q9PC04 PRELIMINARY; PRT: 1190 AA.

```

```

AC Q9PC04.

```

```

DT 01-OCT-2000 (TRENBLREL. 15, Created)

```

```

DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)

```

```

DT 01-OCT-2000 (TRENBLREL. 15, Last annotation update)

```

```

DE SURFACE PROTEIN.

```

```

GN xfl981.

```

```

OS Xylella fastidiosa.

```

```

OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;

```

```

OC Xylella.

```

```

NCBI_TaxID=2371;

```

```

RP SEQUENCE FROM N.A.

```

```

RC STRAIN=9A5C; PubMed=10910347;

```

```

RX MEDLINE=20365717;

```

```

RA Simpson A.J.G., Rehnach F.C., Arruda P., Abreu F.A., Acencio M.,

```

```

RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,

```

```

RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,

```

```

RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,

```

```

RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,

```

```

RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorly H.,

```

```

RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,

```

```

RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,

```

```

RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,

```

```

RA Ho P.L., Hohetsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,

```

```

RA Krieger J.E., Kurame E.E., Laigret F., Lambais M.R., Leite L.C.C.,

```

```

RA Machado E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,

```

```

RA Machado M.A., Madela A.M.B.N., Madela H.M.F., Marino C.L.,

```

```

RA Marques C.F., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,

```

```

RA Menck C.V., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,

```

```

RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Neto L.E.S.,

```

```

RA Nhani A.J.R., Nodrega F.G., Nunes L.R., Oliveira M.A.,

```

```

RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,

```

```

RA Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,

```

```

RA Quaglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,

```

```

RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,

```

```

RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,

```

```

RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,

```

```

RA de Souza A.P., Terenzi M.F., Trufi D., Tsai S.M., Tsuchioka M.H.,

```

```

RA Zago M.A., Zatz M., Meidams J., Setubal J.C.;

```

```

RT "rpe genome sequence of the plant pathogen Xylella fastidiosa.";

```

```

RL Nature 406:151-157(2000).

```

```

DR EMBL; AE004017; AAF84783.1;

```

```

SQ SEQUENCE 1190 AA; 118446 MW; 756741B0C8D787CC CRC64;

```

```

alignment_scores:

```

```

Quality: 168.50 Length: 847

```

```

Ratio: 0.428 Gaps: 40

```

```

Percent Similarity: 46.517 Percent Identity: 21.842

```

```

alignment_block:

```

```

US-09-306-689-12 x Q9PC04 ..

```

```

Align seg 1/1 to: Q9PC04 from: 1 to: 1190

```

```

40 GGCTGCTGCGGCTGAGCGGTTCTCAAGATTGAGCTAGCGCTCGCTCGC 89
      |||:::

```

```

155 GlyLeuValAspLysSerGlyAla.....ThrTyrSerMetArgLys 168
      |||:::

```

```

90 GGGT.....GGCTTAGCCAGCATTTGAGCTAGCGGCC 121
      |||:::

```

```

168 uGlySerIleAlaThrMetAsnGlySerAlaGlnGlyIleAspSerIleAlaI 185
      |||:::

```

```

122 TGGCGCCCTGCGAGCGGTAGCCAA.....GATTGAGCATAGCGGCTG 162

```



```

335 TCCCA..... 339
657 AlarGAlaIaGluleuGlulysGlnlleGlnarGaspGlnIaIale 673
340 .AAATGGGGCAAAAAAATATCTCTATATCCCAAAATACCAAT 388
673 ulysThrGlyGluSer.....ValasnTyr...T 682
389 ATGTACTGAAACAGTAATGTTTACAGATTAGTCAAGCGCGGAA 438
682 ySerThrAlaIaIaGly.....ProValAlaValIylsGluSerLys 695
439 GAGTTGGGAGTTGAGTACAAAGAAAGCAATATATTGCAACAGC 488
696 GluVal.....ValGluAlaArgLeuLysLysasnGlnIuGluTy 709
489 TCACCAAGATTAGGCAAGATTCAACCCGCTATTGCTTAAGTACGCTG 538
709 rGluArgThrThrGlyThrMetAlaLeuGlyAspGlyAlaValAlaLysA 726
539 GCATTGTTATCCGCTCCACAATGTATAATGCTACAGAAACTAAA 588
726 rGluAlaIaLysGluAla.....AlaGluSerGlnlleGluLyslleArg 740
589 GCAGCCCAAGCATTAGTGTCTGCCAAAGCATTTGTACAAATGCAATAA 638
741 AlaspasnglnValPheSerAlaThrPheVal.....Lys 752
639 AGCCAAA...ACTGTATTATCGCATTCATCTATTTTAGGCTCAGTAT 685
752 sAlaArgGlnGluAlaLeuAspLyslleGlnLysIle..... 764
686 TGCTGGAATGATTAGATGAGCCTTACAGATAACAGCAACCAAT 735
765 .....AsnAspAspLysSerLeuSerAspAspGlnLysAsnLys 777
736 GCTTGTGCT.....AAAGCTGGCTTGAGGCTAAC 764
778 LeuLeuAlaProLeuArgGluLysValAsnLysSerTyrLeuGlnIuProAl 794
765 AAATTCATTATGAAATATTTGCTAATTCGTAAAAACACTTGACCAAT 814
794 aGlnLysLeuValTyrAspLeu...SerSerArgLysAsnAlaThrGlu. 809
815 TTGGTGAGCAATATGATCAATTTGCTCAAAATACAAATATCAAGGC 864
810 .....LysGlnlleAlaIaHisPheSerAspMetLeuGlnLysAlaLysLys 824
865 TTAGGACTTTAGGACAAACTCAAAATATTCGCTGACTTGATTAAGC 914
825 GluLysasnThr...GluGlnValGlnLysLeuGlnGly..... 836
915 TGGCTTGTTAGATGTATCTCAGGGCTA.....TTATCGGGCG 955
837 .....SerIleArgGlyTyrGlnGlnHisLeuGlnAlaVal 848
956 CAACAGCTGCATTGTAATTGCTACAGTAAAAATGCTCAACAGTAAAAA 1005
848 aLaIaGlnGlnLeuThrGlnIaIaGlnPheGluArgAspSerAlaIaLys 864
1006 GTGGGTGGGGTTTGAATTGGCAACCA.....GTTGTTGTAATAT 1049
865 ThrGlnLysGlyVal...MetSerAsnGlnGlyThrValLeuGlnI 880
1050 TACC.....AAAGCGTTTCTTCTTACATTTTAGCCC 1081
880 yThrThrAspLysAlaIaGlnLysAlaLeuAlaGlnTyrMetArGasnG 897
1082 AACGTGGGAGGAGTTATCTTCACT..... 1110
897 lmetAspSerAlaIaThrTyrGlnArgThrLeuProAspGlyThAlaMet 913
1111 .....GGGCT..... 1116

```

```

914 LeuAspPheGlnGlyLysProIlelleGlyProLysGlnLeuLysThrGln 930
1116 ..... 1116
930 nLeuAsnLeuGlnLysAlaSerSerAlaSerSerLeuGlnLysMetSerA 947
1117 .....GTGGCTGCTTAAATGCTTCTTACTGTTTCT..... 1146
947 spGlnGluArgAlaIaIaIaIaIaIaLeuThrLysAlaArgGlnGln 963
1147 ...CTTGCATTTAGCCCATTTAGCATTTGCCGATTTGCCATTAATTAA 1193
964 AspAlaIaIaIaIaIaGlnLysAlaGlnArgThrAlaAsnAlaSerGln 980
1194 TCATGCAAAAAGTTTAGAGTTATGCCAAGCTTAAAAAATTAGCT 1243
980 nArgAlaIaIaArgLysGlnGlnAsnAlaGlnArgLysLeuAlaIaGlyT 997
1244 ATGAC.....GGAGATATTTA.....TTAGCAGAA 1269
997 yGlnLysAlaLeuAspLysAlaAspGlnLeuMetGlyGlnMetGlyGln 1013
1270 TATCAGCGGGAGACAGGACTATTGATGATCGGTACTGCAATTAATAC 1319
1014 SerSerLysAlaThrValSerPheAspGlnSerLeuArgAspThrThrLys 1030
1320 CGCATTTGGCGGCTATTCGCTGCTGTG...TCTGCTGCTGACGCGCAT 1366
1030 sSerLeuThrGlnLeuAlaAsnAlaValProAsnGlnPheIleThrGln 1047
1367 TAACATTGAAAAAGTTAAACATTAATCTTCTC..... 1398
1047 lmetIleAspLysAlaLysSerArgLeuAlaAspLeuAlaAsnAlaSer 1063
1399 .....ATCAGCATAGCAAAAAGAGAAAGTACCATTTCA 1433
1064 AspAspTyrArgGlnMetPheAsnArgArgAsnValGlnIuMetIleSe 1080
1434 AAATGCTGCCAGAGGCTGATTTTGTAAAGAGTCCATTAATTAAG 1483
1080 rThrThrAlaProGlnSerAspSerIleIleSerAla...GlyTyrLysP 1096
1484 CACTAAAGATGAGAAATCGAAGAAATCATCGCTCAAAATGGCAGCGG 1533
1096 roSerArgGlnGlnLysValAlaAspPheAsnAspThrTyrAsnArgAsn 1112
1534 ATCAC..... 1539
1113 LeuLysAlaLeuMetAspLeuArgAspGlnAlaSerAspProLysIleVa 1129
1540 .....TCAAGCAAGTGTGATGATTTATGCAAAAGTTAAGCGC. 1578
1129 lAlaLeuTyrThrLysGlnlleAsnGlnLeuValAlaIaGlnLysnThra 1146
1579 .....AAATTAACCAAGATGAGCTATCAAAAGTTGTGATTAAGTAT 1620
1146 lAleuIleLysGlnThrGlyThrAlaThrGlnLysLeuAlaLeuGlnTyr 1162
1621 GAATTGCTCAAAACATFACAAAAATGTGACAAACAGCTTACATAAGTTAAT 1670
1163 GluAsnLeu.....AlaGlnGlnLeuGlnAsnSerTrpSerAsnLeuPh 1177
1671 CTGATCTGTAAGT.....GCATTTA 1690
1177 eSerAsnMetThrAspThrLeuThrAspPheValMetLysGlyLysLeuA 1194
1691 CCTGCTTAATGATTCAGAAATGTATTAAGTCCCAACTTCAATGTG 1740
1194 spPheSerSerLeuAlaGlnSerIleLeuArgAspIleThrAsnMetVal 1210
1741 GATCAAAAGTTATCTTCTTCAATTTGCTAGGGAGATTCAGCATTTGGAG 1790
:::||||::: ::::: |||:::|:::

```

```
1211 .ValylsThrGlnIleThrLeuProLeuMetAsnIleuGlyMetGlyT 1227
1791 CTAGGGCTGGGCC.....CTGGACGGGTCTTC 1819
1227 hTThrAlaIaGlySerSerGlnSerGlyAsnLeuSerGlyValAla 1243
1820 AAGATTGAGCTACGGCTGCGTCCGGGTGGCTTAGCC..... 1858
1244 SerAlaValAlaAsnGlnGlyValArgMetAsnAlaValAsnGlyAsPly 1260
1859 .AGCATTTGA...GCTACGGCCTGGCCCTGGCAGCGGTAGCCAGATTG 1904
1260 sSerValGlyGlnAlaThrLygluThrSerSerSerValSerGlyLeuG 1277
1905 GAGCTACGGCCCTGCGGTGGATCTCAGCATTTGAGCTACGGCCTGC 1954
1277 LygluThrThr...GlnGlnThrThrSerAlaIleGlySerAlaThrAsn 1292
1955 GCCCTG.....GCAGCGGTCTCAGAT 1977
1293 AlaIleGlyAsnTrpValAsnGlyLeuPheThrSerThrGluAlaLysAs 1309
1978 TGGAGCTACGGCCTGCTCCGGGTGGCTTAGCCAGCAT..... 2017
1309 PALaGluThrLysAlaValLysThrSerIlePheSerMetGlnAsnLeuS 1326
2018 ..GGAGCTACGGCCTGCGCCTGGCAGCGGTAGCCAGATTGGAGCT 2062
1326 eRSerValThrGlyAlaLeuSerAlaAlaPheAlaMetIleuGlyAla 1341
```

THIS PAGE BLANK (USPTO)

OM of: US-09-306-689-12 to: SwissProt_39.* out_format : pfs
Date: Mar 2, 2001 10:51 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+np2.model -DBV=rlp
-Q/cgcn2.1/USPRO.spool/US09306689/runat.02032001_102830_9659/app_query.fasta.1.2389
-DB=SwissProt_39 -QPM=fastan -SUFFIX=isp -GAPD=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000
-OGAPD=4.500 -OGAPEXT=0.050 -XGAPD=10.000 -XGAPEXT=0.500
-FCAPD=6.000 -FCAPEXT=7.000 -YGAPD=10.000 -YGAPEXT=0.500
-DELDP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsun62
-TRANS=human0.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFM=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USER=US09306689.@CGN1_1.176 -NCRU=6 -ICPU=3 -LONGLOG -NO_XLPHY
-MAIT -THREADS=1

Search information block:

Query: US-09-306-689-12
Query length: 2088
Database: SwissProt_39.*
Database sequences: 88757
Database length: 32294092
Search time (sec): 97.840000

score_list:

Sequence	Strd Orig	ZScore	Escore	len	Documentation
SwissProt_39:CYAA_BORPE	212.00	136.29	0.0002	1706	P15318 bordetella pertussis
SwissProt_39:CYAA_BORPE	206.50	139.11	0.0003	1705	Q57506 bordetella bronchise
SwissProt_39:TALA_HUMAN	166.71	0.0052	2491	P54633 dictyostella bronchise	
SwissProt_39:K2C1_HUMAN	166.50	0.0394	643	P04264 homo sapiens (human)	
SwissProt_39:K220_HUMAN	166.00	0.0419	638	P01546 homo sapiens (human)	
SwissProt_39:K22E_HUMAN	166.00	0.0419	645	P35508 homo sapiens (human)	
SwissProt_39:K1C1_HUMAN	163.00	0.0603	622	P35527 homo sapiens (human)	
SwissProt_39:GUNE_CELFI	161.50	0.0699	1045	P26225 cellulomonas finl.	
SwissProt_39:PALI_MOUSE	159.00	0.0893	2541	P26039 mus musculus (mouse)	
SwissProt_39:FLIC_SALON	157.50	0.1188	507	Q06974 salmonella oranienber	
SwissProt_39:Y109_YEAST	155.50	0.1449	995	P40642 saccharomyces cerevis	
SwissProt_39:PALI_HUMAN	155.50	0.1363	2541	Q09490 homo sapiens (human)	
SwissProt_39:SM41_HEMPU	154.00	0.1841	407	Q26264 hemiphysalis melanosate	
SwissProt_39:CPN1_DROME	153.00	0.1978	865	Q02910 drosophila melanogaste	
SwissProt_39:FLA3_CAMEL	152.50	0.2158	574	Q46113 campylobacter jejuni.	
SwissProt_39:SLAP_CAMEL	150.50	0.2663	933	P35827 campylobacter fetus.	
SwissProt_39:ANP_NOTCO	145.50	0.4928	790	P24856 notiothenia coriiceps	
SwissProt_39:APMD_PIG	145.00	0.5108	1150	P12021 sus scrofa (pig)	
SwissProt_39:Y116_YEAST	144.50	0.3438	1113	Q00230 saccharomyces cerevi	
SwissProt_39:Y11C1_HUMAN	144.00	0.6019	593	P13645 homo sapiens (human)	
SwissProt_39:SLAP_CAVCR	144.00	0.5808	1025	P35828 caluopacter crescent	
SwissProt_39:MYH_YEAST	143.50	0.6056	1367	P08640 saccharomyces cerevi	
SwissProt_39:RBP1_PYAVB	143.50	0.5771	2869	Q00798 plasmodium vivax (st	
SwissProt_39:FLJB_SALAE	142.00	0.7752	500	P52615 salmonella abortus-eg	
SwissProt_39:SVS2_RAT	141.00	0.8856	414	P22006 rattus norvegicus (rat)	
SwissProt_39:Y190K_RICRI	140.50	0.8427	2249	P15921 rickettsia rickettsi	
SwissProt_39:MCPL_CAVCR	140.00	0.9639	657	Q00986 caluobacter crescent	
SwissProt_39:PLT_PSEAE	140.00	0.9675	682	P42257 pseudomonas aeruginos	
SwissProt_39:FLIC_SALAE	139.50	1.05	507	Q06968 salmonella berta. fld	
SwissProt_39:MYSN_ACGCA	139.00	1.04	1509	P05659 acanthamoeba castelli	
SwissProt_39:G33_BP44	138.50	1.11	1289	P18771 bacteriophage t4. lat	

seq_name: SwissProt_39:CYAA_BORPE

seq_documentation_block:

ID CYAA_BORPE STANDARD; PRT; 1706 AA.
AC P15318;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE BIFUNCTIONAL HEMOLYSIN-ADENYLATE CYCLASE PRECURSOR (CYCLOLYSIN) (ACT)
DE (AC-HLY) [CONTAINS: CALMODULIN-SENSITIVE ADENYLATE CYCLASE
DE (EC 4.6.1.1) (ATP PYROPHOSPHATE-LYASE) (ADENYLATE CYCLASE); HEMOLYSIN].
GN CVA OR CIVA.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
NC [1]
RP SEQUENCE FROM N.A.
RC STRAIN-18323.
RX MEDLINE-68216178; PubMed-2897067;
RA Glaser P., Ladant D., Sezer O., Pichot F., Ullmann A., Danchin A.;
RT "The calmodulin-sensitive adenylyate cyclase of Bordetella pertussis:
RT cloning and expression in Escherichia coli.";
RL Mol. Microbiol. 2:19-30(1988).
RN [2]
RP SEQUENCE OF 1489-1706 FROM N.A., AND BIFUNCTIONAL PROTEIN DESCRIPTION.
RC STRAIN-18323;
RX MEDLINE-69091151; PubMed-2905265;
RA Glaser P., Sakamoto H., Bellalou J., Ullmann A., Danchin A.;
RT "Secretion of cyclolysin, the calmodulin-sensitive adenylyate cyclase-
RT haemolysin bifunctional protein of Bordetella pertussis.";
RL EMBO J. 7:3997-4004(1988).
RN [3]
RP DOMAINS.
RX MEDLINE-91177021; PubMed-2007407;
RA Munier H., Gillies A.-M., Glaser P., Danchin A., Sarfati R., Barzu O.;
RT "Isolation and characterization of catalytic and calmodulin-binding
RT domains of Bordetella pertussis adenylyate cyclase.";
RL Eur. J. Biochem. 196:469-474(1991).
RN [4]
RP MUTAGENESIS.
RX MEDLINE-69251630; PubMed-2542030;
RA Glaser P., Elmaglou-Lazaridou A., Krin E., Ladant D., Barzu O.,
RA Danchin A.;
RT "Identification of residues essential for catalysis and binding of
RT calmodulin in Bordetella pertussis adenylyate cyclase by site-directed
RT mutagenesis.";
RL EMBO J. 8:967-972(1989).
RN [5]
RP MUTAGENESIS.
RX MEDLINE-91266896; PubMed-2050107;
RA Glaser P., Munier H., Gillies A.-M., Krin E., Porumb T., Barzu O.,
RA Sarfati R., Pellequer C., Danchin A.;
RT "Functional consequences of single amino acid substitutions in
RT calmodulin-activated adenylyate cyclase of Bordetella pertussis.";
RL EMBO J. 10:1683-1688(1991).
RN [6]
RP REVIEW.
RX MEDLINE-93119764; PubMed-8418825;
RA Danchin A.;
RT "Phylogeny of adenylyl cyclases.";
RL Adv. Second Messenger Phosphoprotein Res. 27:109-162(1993).
RN [7]
RP PALMITOYLATION AT LYS-983.
RX MEDLINE-95025937; PubMed-7939682;
RA Hackett M., Guo L., Shabanowitz J., Hunt D.F., Hewlett E.L.;
RT "Internal lysine palmitoylation in adenylyate cyclase toxin from
RT Bordetella pertussis.";
RL Science 266:433-435(1994).
RN [8]
RP PALMITOYLATION AT LYS-860.
RX MEDLINE-99214144; PubMed-10196151;
RA Basar T., Havilek V., Bezouskova S., Halada P., Hackett M., Sebo P.;
RT "The conserved lysine 860 in the additional fatty-acylation site of
RT Bordetella pertussis adenylyate cyclase is crucial for toxin function
RT independently of its acylation status.";
RL J. Biol. Chem. 274:10777-10783(1999).
RN [9]
RP FUNCTION: THIS ADENYLATE CYCLASE BELONGS TO A SPECIAL CLASS OF
CC BACTERIAL TOXIN. IT CAUSES WHOOPING COUGH BY ACTING ON MAMMALIAN
CC CELLS BY ELEVATING CAMP-CONCENTRATION AND THUS DISRUPTS NORMAL
CC CELL FUNCTION.
CC -!- CATALYTIC ACTIVITY: ATP = 3',5'-CYCLIC AMP + PYROPHOSPHATE.


```

926 TAGATGTTATCTCAGGCGCTATTATCGGCGCAACGCTGACTTACTT 975
    : : : : : : : : : : : : : : : : : : : : : : : :
539 eAlaLeuGlyGlyGlyLeuAlaAlaAlaValGlyValGlyMetSerLeu 555
    : : : : : : : : : : : : : : : : : : : : : : : :
976 GCAGATTAATAAATGCTTCAACAGCTAAATAAGTGGCGGGTGGTGAATT 1025
    : : : : : : : : : : : : : : : : : : : : : : : :
556 ThrAsp...AspAlaProAlaGlyGlnLysAlaAlaAlaGlyValGlu 571
    : : : : : : : : : : : : : : : : : : : : : : : :
1026 GGCATAACCAAGTTGTTGTAATTAATTAACAAGCCGTTCTTCTTACATT 1075
    : : : : : : : : : : : : : : : : : : : : : : : :
571 eAlaLeuGlnLeuThrGlyGlyThrValGlyLeuAlaSerSerLeuAla 588
    : : : : : : : : : : : : : : : : : : : : : : : :
1076 TA.....GCCCAAGTGTGGCGAGCGATTATCTTCACTGGG 1113
    : : : : : : : : : : : : : : : : : : : : : : : :
588 eAlaLeuAlaAlaAlaArgGlyValThrSerGlyLeuGlnValAlaGly 604
    : : : : : : : : : : : : : : : : : : : : : : : :
1114 CCGTGGCGCTTTAATGCTTCTACTGTTTCTGCGATTAGCCCAATT 1163
    : : : : : : : : : : : : : : : : : : : : : : : :
605 AlaSerAlaGlyAlaAlaAlaGlyAlaLeuAlaAlaLeuSerProMet 621
    : : : : : : : : : : : : : : : : : : : : : : : :
1164 ACCATTGGCGGTATTCGCGATTAATTAATCAATGCAAAAGTTTAGAGA 1213
    : : : : : : : : : : : : : : : : : : : : : : : :
621 tGluLeuTyrGlyLeuValGlnGlnSerHisTyrAlaAspGlnLeuAsp 638
    : : : : : : : : : : : : : : : : : : : : : : : :
1214 GTTATGCCAGACGCTTTAAAAAATTAAGCTATGACGAGATATTATTATTA 1263
    : : : : : : : : : : : : : : : : : : : : : : : :
638 yLeuAlaGlnLeuSerSerAlaTyrGlyTyrGlnGlyLysAlaLeuLeu 654
    : : : : : : : : : : : : : : : : : : : : : : : :
1264 GCAGATTAATTCAGCGGAGACAGGAGTATTAATGATCGTACTGCAATP 1313
    : : : : : : : : : : : : : : : : : : : : : : : :
655 AlaGlnLeuTyrThrAspLysThrAlaAlaGlnGlyAlaValAlaGly 671
    : : : : : : : : : : : : : : : : : : : : : : : :
1314 TAAATACCGCATTTGCCGCTATTCGCTGCTGCTGCT...GCTGCTGCAAG 1360
    : : : : : : : : : : : : : : : : : : : : : : : :
671 lSerAlaValLeuSerThrValGlyAlaAlaValSerIleAlaAlaAla 688
    : : : : : : : : : : : : : : : : : : : : : : : :
1361 CCGATTTAACATTTGAAAAAGTTAAACATTAATCTTGTCAATCAGATTAAC 1410
    : : : : : : : : : : : : : : : : : : : : : : : :
688 lAserValValGlyAlaProVal.....AlaValAlaThrSerLeu 701
    : : : : : : : : : : : : : : : : : : : : : : : :
1411 AAAAAAGACAAAGTGAC.....ATTCAAACTGCTGTCG 1445
    : : : : : : : : : : : : : : : : : : : : : : : :
702 LeuThrGlyAlaLeuAsnGlyLeuAlaArgGlyValGlnGlnProIle 718
    : : : : : : : : : : : : : : : : : : : : : : : :
1446 AGAGGCT.....GATTTGCTAAAGAGTGCCTAAT..... 1476
    : : : : : : : : : : : : : : : : : : : : : : : :
718 eGluLysLeuAlaAsnAspTyrAlaLarGlyLysIleAspGluLeuGly 735
    : : : : : : : : : : : : : : : : : : : : : : : :
1477 .....TATAAGCACTAAAGATGAGAAATCGAAGAAATCATC 1515
    : : : : : : : : : : : : : : : : : : : : : : : :
735 roGlnAlaLarTyrPheGluLysAsnLeuGlnAlaArgHisGlnGlnLeu 751
    : : : : : : : : : : : : : : : : : : : : : : : :
1516 GGTCAAAATGGGAGCGGATCACCTCAAGCAAGATGATGATTTATTCG 1565
    : : : : : : : : : : : : : : : : : : : : : : : :
752 AsnSerAspGlyLeuArg.....LysMetLeuAlaAspLeuAla 765
    : : : : : : : : : : : : : : : : : : : : : : : :
1566 AAAAGTAAAGGCAAA.....ATTACCAAGATGAGCTATCA 1603
    : : : : : : : : : : : : : : : : : : : : : : : :
765 acGlyTrpAsnAlaSerSerValIleGlyValGlnThrThrGlnIleSer 782
    : : : : : : : : : : : : : : : : : : : : : : : :
1604 AAGTGTGTATATCTGAATGCTCAAAACATAGCAAAATGAGCAAAAC 1653
    : : : : : : : : : : : : : : : : : : : : : : : :
782 ys.....SerAlaLeuGlnLeuAlaAlaIleThrGlyLysAlaAsp 796
    : : : : : : : : : : : : : : : : : : : : : : : :
1654 AGCTTAGATAGTTATCTCATCTGTAAGTCATTT 1689
    : : : : : : : : : : : : : : : : : : : : : : : :
797 .....LeuLysSerValaAspValPhe 803
    : : : : : : : : : : : : : : : : : : : : : : : :
seq_name: SwissProt_39:CYAA_BORBR
seq_documentation_block:
ID CYAA_BORBR STANDARD: PRT: 1705 AA.

```

```

AC 057506; 005179;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE BIFUNCTIONAL HEMOLYSIN-ADENYLATE CYCLASE PRECURSOR (CYCLOLYSIN) (ACT)
DE (AC-HLY) [COMTAINS: CALMODULIN-SENSITIVE ADENYLATE CYCLASE
DE (EC 4.6.1.1) (AMP PYROPHOSPHATE-LYASE) (ADENYLATE CYCLASE); HEMOLYSIN].
CN CYA OR CYAA.
OS Bordetella bronchiseptica.
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CIP 9.73;
RX MEDLINE=96009899; PubMed=7557410;
RA Betsou F., Sismelro O., Danchin A., Guiso N.;
RT "Cloning and sequence of the Bordetella bronchiseptica adenylate
RT cyclase-hemolysin-encoding gene: comparison with the Bordetella
RT pertussis gene.";
RL Gene 162:165-166(1995).
RN [2]
RP REVISION TO 1517.
RC STRAIN-CIP 9.73;
RA Danchin A.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS ADENYLATE CYCLASE BELONGS TO A SPECIAL CLASS OF
CC BACTERIAL TOXIN. IT CAUSES WHOOPING COUGH BY ACTING ON MAMMALIAN
CC CELLS BY ELEVATING CAMP-CONCENTRATION AND THUS DISRUPTS NORMAL
CC CELL FUNCTION.
CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-CYCLIC AMP + PYROPHOSPHATE.
CC -1- ENZYME REGULATION: ACTIVATED BY HOST CALMODULIN.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: THE GLY-RICH REGION IS PROBABLY INVOLVED IN BINDING
CC CALCIUM, WHICH IS REQUIRED FOR TARGET CELL-BINDING OR CYTOLYTIC
CC ACTIVITY (BY SIMILARITY).
CC -1- PTM: RELEASED IN A PROCESSED FORM.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO ADENYLATE CYCLASE
CC CLASS-2 FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE RTX
CC PROKARYOTIC TOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL; Z37112; CAAB5481.1; -
DR INTERPRO; IPR001343; -
DR PFM; PF00353; HemolysinCabin; 5.
DR PRINTS; PR00313; CABDNRGPT.
DR PROSITE; PS00330; HEMOLYSIN_CALCIUM; 5.
KW lyase; CAMP synthesis; ATP-binding; Hemolysis; Toxin; Virulence;
KW whooping cough; Calcium-binding; Repeat; Lipoprotein; Palmitate.
FT CHAIN 1 312
FT CALMODULIN-SENSITIVE ADENYLATE
FT CYCLASE.
FT HEMOLYSIN (BY SIMILARITY TO E. COLI
FT HEMOLYSIN HLYA).
FT A. CATALYTIC.
FT B. ALA/GLY-RICH.
FT C.
FT D. ASP/GLY-RICH.
FT ATP (POTENTIAL).
FT 28 X REPEATS, GLY-RICH.
FT REPEAT 1014 1019 1.
FT REPEAT 1023 1028 2.
FT REPEAT 1032 1037 3.
FT REPEAT 1041 1046 4.
FT REPEAT 1050 1055 5.
FT REPEAT 1059 1064 6.

```

```

FT REPEAT 1079 1084 7.
FT REPEAT 1164 1169 8.
FT REPEAT 1173 1178 9.
FT REPEAT 1182 1187 10.
FT REPEAT 1202 1207 11.
FT REPEAT 1279 1284 12.
FT REPEAT 1288 1293 13.
FT REPEAT 1297 1302 14.
FT REPEAT 1306 1311 15.
FT REPEAT 1315 1320 16.
FT REPEAT 1324 1329 17.
FT REPEAT 1344 1349 18.
FT REPEAT 1420 1425 19.
FT REPEAT 1429 1434 20.
FT REPEAT 1438 1443 21.
FT REPEAT 1447 1452 22.
FT REPEAT 1452 1457 23.
FT REPEAT 1555 1560 24.
FT REPEAT 1564 1569 25.
FT REPEAT 1573 1578 26.
FT REPEAT 1582 1587 27.
FT REPEAT 1592 1597 28.
FT REPEAT 1604 1609 29.
FT LIPID 859 864 30.
FT LIPID 982 987 31.
SQ SEQUENCE 1705 AA: 177249 MW: D8530697A4BAE60E CRC64;

```

```

alignment_scores:
  Quality: 206.50      Length: 372
  Ratio: 1.059         Gaps: 11
  Percent Similarity: 52.419   Percent Identity: 23.387

```

alignment_block:

us-09-306-689-12 x CYAA_BORBR ..

Align seg 1/1 to: CYAA_BORBR from: 1 to: 1705

```

300 TTGGAGCTAGCGCTGCGTCCGGGTGATCTAGCTCCCAAAAACGTGGG 349
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
380 LeuGlnIuValProAlaSer.ProGlyLeuArgArgProSerLeuVal 396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
350 CAAAAAATATTCCTATATTCCTATATTCCTATATTCCTATATTCCTAT 399
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
396 IValGlu.....ArgGlnAspSerGlyTyrAspSerLeu 407
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
400 CAAGGTATGTTTACAGATTAGTCAAGGCGCGAGAGTGGGGAT 449
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
408 AspGlyValGlySerArgSerPhe..... 415
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
450 TGAAGTACAAAGAGAGAACGCAATTAATTCGAACAGCTCAACAGTT 499
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
416 .....SerL 417
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
500 TAGGACGATTCAACCGCTATTGGCTTAAGCTAGCGTGGCATGTGTTA 549
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
417 euGlyGluValSerAspMet.....AlaIVal 426
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
550 TCCGCTCCACAATTTGATTAATTCCTACAGAAAACCTAAGAGCGCAAGC 599
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
427 GluAlaIValLeuGluMetThrArgGlnValLeuHisIValGlyAla 443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
600 ATTAGTCTCCGCAAGATTTGACAAAATGCAATTA.....GCCA 643
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
443 gGlnAspAlaIValProGlyValSerGlyAlaSerAlaHisIValGly 460
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
644 AAAGCTATATCTGGCATCTATTTTGGCTCAGATTGGCTGGA 693
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
460 IValGlnAlaIValGlnGlyAlaIValAlaIValAlaIValAlaIVal 472
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
694 ATGATTTAGATGAGGCTTACAGATTAACAGAACCAACATGCTCTTGC 743
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
473 .....GlnArgLeuValHisAlaIVal 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

744 T.....AAAGCTGCTTGGAGCTAACAAATTCATTAA 775
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
480 aleuMetThrGlnPheGlyArgIValGlySerThrAsnThrProGlnIVal 497
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
776 TTGAAATATTCCTAATTCAGTAAACACTTGCAGATTTGGTGACAA 825
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
497 IValSerLeuSerAlaIValIValPheGlyLeuGlyGluAlaSerSerAla 513
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
826 ATTAGTCAA.....TTGGTTCAAACTACAAATATTCAA 860
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
514 ValAlaIValThrValSerGlyPheArgGlySerSerArgTyrIValGly 530
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
861 AGCGTTAGGACTTTAGAGACAACCTCAAAAATATGCTGGAGCTTGATA 910
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
530 YGlyPheGlyValAlaGlyGlyAlaMetAlaLeuGlyGlyGlyIValGly 547
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
911 AAGCTGGCTTGGTTAGATGTTATCTCAGGCTATATATGCGGCGACAA 960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
547 IValGlyAlaGlyMetSer..... 553
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
961 GCTGCACTTGTACTTGCAGATTAATAATGCTTCAACAGCTAAAGTGGG 1010
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
554 .....LeuThrAsp..AspAlaProAlaGlyGlnIValAlaI 565
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1011 TCGCGCTTTGAAATTCGCAACCAAGTGTGTGTAATTAATTCACCAAGCC 1060
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
565 aAlaGlyAlaGlnIValLeuAlaLeuGlnIValGlyThrValGlnIVal 582
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1061 TTTCTTCTTACATTTA.....GCCACAGCTGTTCAGCAGGT 1098
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
582 IValSerIValLeuAlaIValAlaIValAlaIValArgIValThrSerGly 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1099 TTATCTCAACTGGGCTGTGGCTGTTTAAATGCTTACTGCTGTTCTCT 1148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
599 LeuGlnIValAlaGlyAlaSerAlaIValAlaIValAlaIValAlaIVal 615
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1149 TGGGATTAGCCCATTTAGCATTTGGCGGTATTCGCGATTAATTAATCAG 1198
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
615 aAlaLeuSerProMetGlnIValGlyGlyLeuValGlnGlnSerHisIVal 632
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1199 CAAAAATTTAGAGACTTATTCGCAACGCTTAAATAATTAAGCTATGAC 1248
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
632 IValGlnIValLeuAspIValLeuAlaGlnIValSerAlaIValGlyGly 648
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1249 GGAGATTAATTTATTCAGCAATATACAGCGGGAACAGGACTATTCATGC 1298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
649 GlnAspAlaLeuLeuAlaGlnIValThrArgAspIValThrAlaIValGly 665
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1299 ATCGGTTACTGCAATTATACCGCATTTGCCGCTATTCGTGTGTGT 1348
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
665 YAlaIValAlaGlyValSerAlaIValLeuSerThrValGlyAlaIVal 682
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1349 CTGCTGCTGCAGCC 1362
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
682 erIValAlaIValAla 686
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

seq_name: SwissProt_39:TALA_DICDI

seq_documentation_block:

```

ID TALA_DICDI STANDARD; PRT; 2491 AA.
AC P54633;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FILOPODIN (TALIN HOMOLOG).
GN TALA.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Dictyostelida; Dictyostelium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX2.
RX MEDLINE=95213284; PubMed=7698984;
RA Kreilmeier M., Gerisch G., Heizer C., Mueller-Taubenberger A.;

```

"A talin homologue of Dictyostelium rapidly assembles at the leading edge of cells in response to chemoattractant.";
 Rt Cell Biol. 129:179-188(1995).
 CC -1- FUNCTION: ACTIN-BINDING PROTEIN THAT MAY BE INVOLVED IN THE
 CC CONTROL OF CELL MOTILITY AND CHEMOTAXIS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; RAPIDLY ASSEMBLES AT THE
 CC LEADING EDGE OF CELLS IN RESPONSE TO CHEMOATTRACTANT.
 CC -1- SIMILARITY: CONTAINS A DOMAIN FOUND IN BAND 4.1, EZRIN, MOESIN,
 CC RADIXIN, AND TALIN.
 CC -1- SIMILARITY: IN THE C-TERMINUS TO YEAST SLA2 AND C.ELEGANS ZK370.3.
 CC PARTIAL, TO YEAST ROD1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U14576; AAC46586.1; -
 DR DICTYDB: DD01121; TALA.
 DR INTERPRO: IPR00299; -
 DR INTERPRO: IPR002558; -
 DR PFAM: PF00373; Band_41; 1.
 DR PFAM: PF01608; L.WEO; 1.
 DR PROSITE: PS00660; BAND_41_1; 1.
 DR PROSITE: PS00661; BAND_41_2; 1.
 DR PROSITE: PS50057; BAND_41_3; 1.
 DR Structural protein; Cytoskeleton; Actin-binding.
 FT DOMAIN 134 294 BAND 4.1-LIKE.
 SEQUENCE 2491 AA; 26881 MW; B24FF0224F2A419B CRC64;

alignment_scores:
 Quality: 182.50 Length: 650
 Ratio: 0.583 Gaps: 31
 Percent Similarity: 48.154 Percent Identity: 22.000

alignment_block:

US-09-306-689-12 x TALA_DICDI ..

Align seg 1/1 to: TALA_DICDI from: 1 to: 2491

```

412 TTACAGATTAGTCAAGCGCCGAGAGTTGGGATTGAGTACAAAG 461
    ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1064 LeuGlyAspLeuValSerAlaThrGluSerAlaSerValAla11 1080
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
462 AGAAGAACCAATTAATTTGCAACAGCTCAACAGCTTAGCAGCATTC 511
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1080 eSerAlaLysThrLeuIleCysAla.....ThrThrGlyLysGlnValG 1095
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
512 AAACCGCTATT...GGCTTAAGTACAGCTGCGATTGTTATCCGCTCA 558
    |||| :: :: :: :: :: :: :: :: :: :: :: :: ::
1095 LnthrLysLeuMetGlyIleThrLysGlnLeuMetIle..... 1107
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
559 CAAATGATGAATTTGCTACAGAAACTAAAGCAGCCCAAGCATTAAGTTC 608
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1108 AspMetGlnGlnLeuIleArg.....Al 1115
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
609 TGCAGCAACATTTGTACAAATGCAATTAAGCCAAAGCATTAATTA... 654
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1115 aSerAlaGSerValArgSerAlaProAsnAspArgArgSerGlnLeuLeu 1132
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
655 .....TCTGCATTCAATTAATTTTGGCTACAGTATTTGGCTGGA 693
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1132 euAspArgArgSerAsnAspValSerIleSerThrAlaIleValGly 1148
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
694 .....ATGATTTAGATGAGCGCTTACAGAAATA 722
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
1149 SerThrAlaAsnValAspCysLysGlnIleAlaSerAlaSerAla 1165
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
723 GCGAACCAACATGCTCTTGTAAAGCTGCGTTGAGCTAACAATTA 771

```

```

1165 eSerAlaLeuLeuSerLeuLysMetGlySerLeuGluSerIleLeuSerG 1182
    ||||| ::||| ::::::||| |||
1182 LnthrThrGlnGluPheAlaPheTyrValGlnGluIleAlaSerSerThr 1198
    ::||| ::||| ::||| ::||| ::|||
799 AAACACTTACCAATTTGGTGAAGCAAAATTAATTTGCTTCAAACT 848
    ::||| ::||| ::||| ::||| ::|||
1199 LysAlaLeuAsnAlaAlaSerGlnGlnValValAlaMet...AlaArgAs 1214
    ::||| ::||| ::||| ::||| ::|||
849 ACAAAATATCAAGCGCTTAGGACTTTAGAGACAAACAAATTAATTCG 898
    ::||| ::||| ::||| ::||| ::|||
1214 nLysAsnLeuLysGlnGlnValAla..... 1222
    ::||| ::||| ::||| ::||| ::|||
899 GTGAGCTTGAATTAAGCTGCGCTTGTAGATGTATTCAGGCGTATTA 948
    ::||| ::||| ::||| ::||| ::|||
1223 .....SerAlaLysIleThrAlaSerAlaLeuSerThrLeuVal 1235
    ::||| ::||| ::||| ::||| ::|||
949 TCGGGCGCAACAGCTGCATTTGACTT.....GCAGATATAAA 986
    ||| ||| ::||| ::||| ||| ||||
1236 SerHisAlaGlnAsnAlaIleValLeuThrGluAsnGlnValaThrLysAs 1252
    ::||| ::||| ::||| ::||| ::|||
987 TGCTTCAACAGCTAAATAAGTGGGCGGCTTGTGAATTTGCCAACCAG 1036
    |||| ::||| ::||| ::||| ::|||
1252 nAlaIleLeuAlaSerThrValAlaLeuGly.....GlyGlnI 1265
    ::||| ::||| ::||| ::||| ::|||
1037 TTTGTTGCT..... 1044
    ::||| ::||| ::||| ::||| ::|||
1265 IetIleGlyLeuLeuAspPheSerLysAlaArgIleAlaAsnTyrLysAsp 1281
    ::||| ::||| ::||| ::||| ::|||
1045 .....AATATTACCAAGCGCTTCTCTTA 1070
    ||| ::||| ::||| ::||| ::|||
1282 ProIleTyrAspGlnAsnLeuIleAsnGlnAlaLysSerValGlnAspH 1298
    ::||| ::||| ::||| ::||| ::|||
1071 CATTTAGGCCCAAGCGTGGCAGCAGCTTATCTCAACGCG..... 1113
    ::||| ::||| ::||| ::||| ::|||
1298 SLeuVal.....LysValGlyArgSerLeuGlyGlyAspGlyAsnAsn 1313
    ::||| ::||| ::||| ::||| ::|||
1114 .....CCGTGGCTGCTTAAATGCTTCTACTGTTCTCTT 1149
    ||| ||| ::||| ::||| ::|||
1313 hTrIleCysAspGlnAlaValAlaAspArgIleIleGlnAlaThrArgSerLeu 1329
    ::||| ::||| ::||| ::||| ::|||
1150 .....GCGATTAGCCCATTTGCCGCTATTTGCCGATTAATTA 1193
    ::||| ::||| ::||| ::||| ::|||
1330 AspLysThrIleLeuPro...AspThrSerGlyLeuGlnThrAsnAlaH 1345
    ::||| ::||| ::||| ::||| ::|||
1194 T.....CATGCAAAAGTTTAGAGCTTATGCCGAGCGCTTA 1231
    ::||| ::||| ::||| ::||| ::|||
1345 SLeuGlnMetLeuHisGlnIleSerLeuLeuAlaIleThrGlnAlaSerL 1362
    ::||| ::||| ::||| ::||| ::|||
1232 AAAAATTTAGGCTATGACGAGATTAATTTATGACAGAAATTCAGCG... 1278
    ||||| ::||| ::||| ::||| ::|||
1362 ystLysLeuGlySerIleThrSerAsnLeuValAsnSerLysAsnAsnSer 1378
    ::||| ::||| ::||| ::||| ::|||
1279 .....GGAACAGGACTATTGATGCA..... 1299
    ::||| ::||| ::||| ::||| ::|||
1379 AspLeuValGlySerGlySerThrAspAlaGlnArgIleIleGlnMetI 1395
    ::||| ::||| ::||| ::||| ::|||
1300 .....TCCGTTACTGCAATTAATACCG 1321
    ::||| ::||| ::||| ::||| ::|||
1395 eGlnAlaIleAlaLysHisValAlaHisCysSerIleSerThrTyrAsnPro 1412
    ::||| ::||| ::||| ::||| ::|||
1322 CATTTGCCCGCTATTTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1371
    ::||| ::||| ::||| ::||| ::|||
1412 splLeuLeuProAlaLysSerIleLeuAspAlaSerGlnMetLeuThr 1428
    ::||| ::||| ::||| ::||| ::|||
1372 TTTGAAAAAGTTTAACATTAATTTGATCATGACGATTAAGCAAAAGAGA 1421
    ::||| ::||| ::||| ::||| ::|||
1429 AlaAsnGlnAlaAspValAsnHisValLeuSerHisAla..... 1441
    ::||| ::||| ::||| ::||| ::|||
1422 AGTGACCATTTCAAAACTGTTCCGAGAGCGTGAATTTTGGTAAAGAGTGC 1471

```


CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: AF237621: AAF60327.1: -
 DR EMBL: M98776: AAB47721.1: -
 DR PIR: A02950: KRH02.
 DR PIR: A22940: A22940.
 DR HSSP: P05412: 1FOS.
 DR AARHU/GHENT-2DPAGE: 4606; NEPHE.
 DR MIM: 139350; -
 DR MIM: 113800; -
 DR INTERPRO: IPR001664; -
 DR INTERPRO: IPR003054; -
 DR PFAM: PF00038: filament; 1.
 DR PRINTS: PR01276: TYPE2KERATIN.
 DR PROSITE: PS00226: IF; 1.
 KM Intermediate filament; Coiled coil; Heptad repeat pattern; Keratin;
 KM Disease mutation; Polymorphism; Phosphorylation.
 FT INIT_MET 0 0
 FT DOMAIN 1 178 HEAD.
 FT DOMAIN 179 488 ROD.
 FT DOMAIN 489 643 TAIL.
 FT DOMAIN 179 214 COIL 1A.
 FT DOMAIN 215 233 LINKER 1.
 FT DOMAIN 234 325 COIL 1B.
 FT DOMAIN 326 349 LINKER 12.
 FT DOMAIN 350 488 COIL 2.
 FT SITE 432 432 STUTTER.
 FT DOMAIN 1 150 GLY/PHE/SER-RICH.
 FT DOMAIN 501 640 GLY/SER-RICH.
 FT MOD_RES 65 65 PHOSPHORYLATION (BY SIMILARITY).
 FT VARIANT 154 154 V -> G (IN EHK).
 FT VARIANT 160 160 L -> P (IN EHK).
 FT VARIANT 185 185 S -> P (IN EHK).
 FT VARIANT 187 187 N -> S (IN EHK).
 FT VARIANT 192 192 S -> P (IN EHK).
 FT VARIANT 311 311 I -> V.
 FT VARIANT 329 329 I -> T.
 FT VARIANT 357 357 N -> Y.
 FT VARIANT 489 489 E -> Q (IN EHK).
 FT VARIANT 536 536 G -> C.
 FT VARIANT 632 632 R -> K.
 FT VARIANT 559 555 /FTID=VAR_003863.
 FT MISSING (IN ALLELE 1B).
 FT /FTID=VAR_003864.
 SQ SEQUENCE 643 AA: 65886 MM: DF945DC462257850 CXC64:

alignment_scores:
 Quality: 166.50 Length: 753
 Ratio: 0.483 Gaps: 31
 Percent Similarity: 45.817 Percent Identity: 20.717

alignment_block:
 US-09-306-689-12 x K2C1_HUMAN ..
 Align seg 1/1 to: K2C1_HUMAN from: 1 to: 643

40 GGCTGGCCCTGGCAGCGCTTCTCAAGATTGAGCTAGCGCTG..... 84
 9 GLYTYRARGSERGLYGLYPHESESERGLYSERLALGLILEALS 25
 85CGTCCGGGTGGCTCA 100
 25 nTYGLNARGARGThrThrSerserSerThrARGserGLYGLY 42
 101 GCCACGATTGG.....AGTACGGCCGCTGCGC 126
 42 LYLGLNARGPHESESERGLYGLYGLYGLYSERPHGLYALGLY 58
 127 CCGGAGCGCTGAGCCCAAGATTGAGCTAGCGCTCCGCGGTGATC 176
 59 GLYGLYPHEGLYSERARGSERLEU.....VALASNLEULYGLYSE 72
 177 TCAGCATTTGAGCTAGCGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGA 226
 72 RLYSSERLIESERLIESERVALALARGLYGLYARGGLYSERCLYP 89
 227 GCTAGGCGCTGCGTCCGGGTGCTAGCCAGCATTTGAGCTAGCGCTG 276
 89 HEGLYGLYGLYTYRGLYGLYGLYPHEGLYGLYGLYPHEGLY... 104
 277 CGCCCTGGCAGCGCTAGCCCAAGATTGAGCTAC..... 309
 105GLYGLYGLYPHEGLYGLYGLYGLYGLYGLYGLYGLYGLY 119
 310 .GGCTGCGCGCGGTGATCTAGCTTC..... 336
 119 YGLYPHEGLYSERGLYGLYGLYPHEGLYGLYGLYGLYPHEGLYGLY 136
 337CCAAAACCTGGGCAAAA 354
 136 LYLGLYTYRGLYGLYTYRGLYPROVALCYSERPROGLYGLYILEGLN 152
 355 AANAATATC.....CTCTATTCCCAAAATATTACCAATATGA 392
 153 GLUVALThrILEASNINserLEUGLInProLEUASNValGLILEALS 169
 393 TACTGAACAAGTAATAGTTTACAGGATTATTCAAAGCGCCGAGAGT 442
 169 PROGLU.....ILEGLN..... 173
 443 TGGGATTTAGGTACAAAGAGAAAGCAATATATTCCAAAGCTCAA 492
 174LYSVALYSSERARGGLUARGGLUINLILEYSSERLEUASN 187
 493 ACCAGTTTAGGCACG...ATTCAACCGCTATTGGCTTACTGAGCGGTG 539
 188 ASNGlnPHEALserPHEILEASPLYSValARGPHELEGLUINLINS 204
 540 CATTGTATTCCGCTCCACAATTTGATTAATGCTACGAAA..... 582
 204 nGLINValLEGLNThrLYSTR.....GLULEUGLInGLINValASP 219
 583ACTAAGCGAGCCAGCATTTAGGTTCCGCGAAGCATTTGCAA 627
 219 hSERThrARGThrHISASNLEUGLInProTYRPhELISERPhELEASN 235
 628 AATGCAATAAAGCCAAAACGTGATTATCTGGCATTCATCTATTATTAG 677
 236 ASNLEUARGARGValASPGLINLEULYSERASPGLINserARGLEUAS 252
 678 CTCAGTATTGGCTGGAATGATTATAGATGAGGCTTACAGATAACAGCA 727
 252 pSERGLULEULYSASNmet..... 258
 728 ACCAACATGCTCTTGTAAAGCTGCGTTGAGCTAACAAATTCATTAAAT 777
 259GLINSPMETVal 262


```

1332 TATTGCTGCTGCTGCTGCTGCTGCGAGCCGATTAACTTGAAGAAAAG 1381
    ::::
410 gleeu..... 411
1382 TTAACATATCTTGTCTATCAGCAATACGAAAAAGAGAGTGCATT 1431
    ::::
412 .....Arglaaglullegluasnvallyslgslasnalaasnleu 425
1432 CAAAGCTGTCGAGAGCGGTGATTGCTTAAGAGTGCCTAATTATA 1481
    ::::
426 Glnthralailealaglualaglulnargllyglumec..... 438
1482 AGCAACTAAGATGAGAAATCGAAGAAATCATGCTCAAAATGCGAGC 1531
    ::::
439 .AlaLeulysasp..... 442
1532 GGATCACCCTCAAGCAAGCTGATGATCTT.....ATGCAAAAGCT 1572
    ::::
443 .....AlaasnalaalysleuGlnaspLeuGlnthralAlaLeuGlnLys... 456
1573 AAGGCAAAATATACCAAGATGAGCTATCAAAAGTGTGTAATCTATGA 1622
    ::::
457 .....AlaLysASPaspLeuAlaArgLeuLeuAlaArgAspTYRGI 469
1623 ATTGCTCAAAATAGCAAAATGTGACAAACAGCTTAGAT..... 1662
    ::::
469 ngluleu.....MetAsnVallyslLeuAlaLeuaspAlaGluileA 483
1663 .....AAGTAACTCATCTATCTGAATGCAATTAATCTGCTCATAT 1701
    ::::
483 lathrTYRArglyslLeuLeuGlnGlyGlnGlyArgmeSerlyglu 499
1702 GATTCGAGAAATGATTAGTGTCTCAACTCAATGTTGATCAAGTTT 1751
    ::::
500 CysGlnSerAlaValCysIleSerValValSer.....AsnVal 512
1752 ATCTTCTCTTCAATTTGCTAGGAGATCTCAGCATTTGAGACGCGCTGC 1801
    ::::
512 lThSerThrserySerlySerlySerSerArglyValphelGlyly 529
1802 GCCCTGCGAGCGGTTCTCAAGATTGAGCTACGCGCTCGCGGTCGC 1851
    ::::
529 alSerGlySerlySer.....GlyGlyTYRlyslGly 540
1852 TCTAGCCGAGCTTGGAGC.....TAGCGCTCGCGCGCTGCGAGCGGTG 1895
    ::::
541 SerSerSerSerSerSerlyTYRlyGlyValSerGlylySerGlySe 557
1896 CCAAGATTGAGCTACGCGCTGCTCGCGGTGATCTCAGCATTTGAGCT 1945
    ::::
557 r.....GlyTYR.....GlyGlyValSerSerlySert 567
1946 ACGGCTGCGCGCTGCGAGCGGTCT...CAAGATTGAGCTACGCGCTG 1992
    ::::
567 hrglyGlyArglySerlySerlySerlyGlnSerSerSerSerlySer 583
1993 CGTCGCGGTGC.....TCTAGCCGAGCTTGGAGCTACGCGCT 2030
    ::::
584 ArgLeuGlyGlyAlaGlySerIleSerValSerIleSerGlyMetGlySe 600
2031 GCGGCTGCGAGCGGTGAGCCAAAGTTGAGCTACGCGCTGCGCGGTG 2080
    ::::
600 rSerSerGlySerIleGlnThrseryGlySerGlyTYRlySerSerlyG 617
2081 GA 2082
    ::::
617 ly 617

```

seq_name: SwissProt_39:K22E_HUMAN

seq_documentation_block: ID K22E_HUMAN STANDARD: PRT: 645 AA.

```

AC P35908;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KERATIN, TYPE II CYTOSKELETAL 2 EPIDERMAL (CYTOKERATIN 2E) (K2E)
DE (CK 2E).
GN KRT2A OR KRT2E.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RC SEQUENCE FROM N.A.
RC TISSUE=HIGH EPIDERMIS;
RX MEDLINE=92380238; PubMed=1380918;
RA Collin C., Moll R., Rubicka S., Ouhayoun J.-P., Franke W.W.;
RT "Characterization of human cytokeratin 2, an epidermal cytoskeletal
RT protein synthesized late during differentiation."
RL Exp. Cell Res. 202:132-141(1992).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS IBS TYR-192 AND LYS-482.
RA Smith F.J.D., Maling C., Covello S.P., Higgins C., Schmidt M.,
RA Lane E.B., Uitto J., Leigh I.M., McLean W.H.I.;
RT "Genomic organization and fine mapping of the keratin 2e gene (KRT2E):
RT kze VI domain polymorphism and novel mutations in ichthyosis bullosa
RT of Siemens."
RL J. Invest. Dermatol. 111:817-821(1998).
RN [3]
RP VARIANT IBS LYS-493.
RX MEDLINE=94358459; PubMed=7521371;
RA McLean W.H.I., Morley S.M., Lane E.B., Eady R.A., Griffiths W.A.,
RA Pajlg D.G., Harper J.I., Higgins C., Leigh I.M.;
RT "Ichthyosis bullosa of Siemens -- a disease involving keratin 2e."
RL J. Invest. Dermatol. 103:277-281(1994).
RN [4]
RP VARIANTS IBS PRO-187 AND LYS-493.
RX MEDLINE=94358461; PubMed=8077693;
RA Kremer H., Zeeuwen P., McLean W.H.I., Mariman E.C.M., Lane E.B.,
RA van de Kerkhof P.C.M., Ropers H.-H., Steijlen P.M.;
RT "Ichthyosis bullosa of Siemens is caused by mutations in the keratin
RT 2e gene."
RL J. Invest. Dermatol. 103:286-289(1994).
RN [5]
RP VARIANTS IBS ASP-493 AND LYS-493.
RX MEDLINE=95038833; PubMed=7524919;
RA Rothnagel J.A., Traupe H., Wojcik S., Huber M., Hohl D.,
RA Pittelkow M.R., Saeki H., Ishibashi Y., Roop D.R.;
RT "Mutations in the rod domain of keratin 2e in patients with ichthyosis
RT bullosa of Siemens."
RL Nat. Genet. 7:485-490(1994).
RN [6]
RP VARIANT IBS PRO-485.
RA Yang J.-M., Lee S., Bang H.-D., Kim W.-S., Lee E.-S., Steinert P.M.;
RT "A novel threonine-to-proline mutation at the end of 2B rod domain in
RT the keratin 2e chain in ichthyosis bullosa of Siemens."
RL J. Invest. Dermatol. 109:116-118(1997).
RN [7]
RC -1- FUNCTION: PROBABLY CONTRIBUTES TO TERMINAL CORNIFICATION.
CC -1- SUBUNIT: HETEROTRIMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC -1- TISSUE SPECIFICITY: IN THE SUPRABASAL LAYERS OF EPIDERMAL
CC TISSUES FROM MOST BODY SITES, EXCEPT IN FORESKIN, IN SQUAMOUS
CC METAPLASIAS AND CARCINOMAS.
CC -1- DEVELOPMENTAL STAGE: SYNTHESIZED DURING MATURATION OF EPIDERMAL
CC KERATINOCYTES.
CC -1- DISEASE: DEFECTS IN KERATIN 2E ARE ASSOCIATED WITH ICHTHYOSIS
CC BULLOSA OF SIEMENS, A RARE AUTOSOMAL DOMINANT DISORDER DISPLAYING
CC A TYPE OF EPIDERMOLYTIC HYPERKERATOSIS CHARACTERIZED BY EXTENSIVE
CC BLISTERING FROM BIRTH. HYPERKERATOSIS AND SHEDDING OF THE OUTER
CC LAYERS OF THE EPIDERMIS (MOLTING) ARE OBSERVED IN LATER WEEKS.
CC -1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
CC MICROFIBILLAR KERATIN: I (ACIDIC: 40-55 KDA) [K9 TO K20] AND II
CC (NEUTRAL TO BASIC: 56-70 KDA) [K1 TO K8].
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC
CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION

```



```

368  rgserlysgluhlaialetyrlhsrlyss..... 380
1216  TATGCCAAGCCTTAAATAATTAGCATGACGAGATATTATTAGC 1265
381  TTTCTGluTleuGlnValThValGlyArgHisGlyAspSerLeu.... 395
1266  AGAATATCAGCGGGAACAGGAGCTATTGATGATCGTTTACTGCAATTA 1315
396  .....lysGluIleIysIleGluIleAspValLeuA 406
1316  ATACCGCATGGCGGCTATGCTGGTGTGTCTGCTGTCAGCCGAT 1365
406  snarValIleGlnArgLeuGlnGly..... 414
1366  TTAACATTGAAAAAGTTAAACATATCTGTCTACGATACGAAATAA 1415
415  .....GluIleAlaHisValLeuLys 421
1416  AGAAGAAAGTGCATTCAAACCTGTTCCGAGAGCTGATTTTCTTAAG 1465
421  sGlnCylAspValGlnAspAlaIleAlaAspAlaGlnArgGlyG 438
1466  AAGTGCCTAATTATTAAGCAACTAAAGATGAGAAA..... 1500
438  Lu.....HisAlaLeuLysAspAlaArgAsnLysLeuAsnAsp 450
1501  ATCAAGAAATCATCGGCAAAATGGCGAGCGGATCACCCTCAAGCAAGT 1550
451  LeuGlnGluAlaLeuGlnGln..... 457
1551  TGATGATCTTATCGCAAAAGTAAAGCAAAATTAACCAAGATGAGTAT 1600
458  .....AlaLysIleAspLeuA 463
1601  CAAAAGTTGTGATTAATGATGATGCTCAAAACATAGCAAAATGTGACA 1650
463  laArgLeuArgAspTyrGlnGlnLeu.....MetAsnValLys 476
1651  AACGCTTAGT.....AAGTTAATCTCATCTGT 1679
477  LeuAlaLeuAspValGluIleAlaThTyrArgLysLeuGlnGlyG 1693
1680  AAGTCAATTTACTGCTCTAATGATGATGAGAAATGATTAAGTGCCTCAA 1729
493  uGlnCylArgMetSerGlyAspLeuSerSerAsnValThrAlaSerValT 510
1730  CTTCATGTTGGATCAAGATTATCTTCTTCAATTGCTAGGGGATCT 1779
510  hrSer.....SerThrIleSerSerAsnValAlaSerLysAla... 522
1780  CAGATTTGAGACTACGGCTCGCCGCGGCGGCTTCTCAAGATTTGAG 1829
523  .....AlaPheGlyLysSerGlyValArgLysSer.....Se 534
1830  CTACGCGCTGCTCGGCTGCTAGCAGCATTTGAGTACGAGCCCTGC 1879
534  rGlyGlyGlyTyrSerSerGlySerSerSerTyrGlySerGlyLysArg 551
1880  GCCCTGGACGGTAGCCAGATTTGAGTACGAGCCCTGCTCGGAGTGA 1929
551  InSerGlySerArgGlyLysSerGlyGlyGlySerIleSerGlyGly 567
1930  TCTGAGATTGAGCTACGGCTCGCCCTGCGGCTTCTCAAGATTTG 1979
568  .....GlyTyrGlySerGlyGlyLysSerGlyLysArgTyrG 580
1980  GAGCTACGCGCTGCTCGGCTGCTGCT..... 2007
580  ySerGlyGlyLysSerLysGlyLysSerIleSerGlyGlyLysTyrG 597
2008  .....AGCCAGCATTTGAGTACGCGCTCGCCCTGCGAGCGGATAGC 2049

```

```

597  erglyGlyLysHisSerSerGlyLysSerArgGlyLysSer 613
2050  CAAGATTGAGACTACGCGCTCGGCTGAGTATCC 2085
614  SerGlyGlyLysSerGlyLysSerGlyLysSer 625
seq_name: SwissProt_39:K1C1_HUMAN
seq_documentation_block:
ID      K1C1_HUMAN          STANDARD:      PRT:      622 AA.
AC      P35527, Q14665.
DT      01-JUN-1994 (Rel. 29, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DE      15-JUL-1998 (Rel. 36, Last annotation update)
DE      KERATIN, TYPE I CYTOSKELETAL 9 (CYTOKERATIN 9) (K9) (CK 9).
GN      KRT9.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
RN      [1]
RP      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC      TISSUE=FOOT SOLE TISSUE;
RX      MEDLINE=94131202; PubMed=7507869;
RA      Langbein L., Heid H.W., Moll I., Franke W.W.;
RT      "Molecular characterization of the body site-specific human epidermal
RT      cytokeratin 9: cDNA cloning, amino acid sequence, and tissue
RT      specificity of gene expression.";
RL      Differentiation 55:57-72(1993).
RN      [2]
RP      SEQUENCE OF 449-465.
RX      MEDLINE=90267446; PubMed=2140676;
RA      Rosen E.M., Meromsky L., Romero R., Setter E., Goldberg I.;
RT      "Human placenta contains an epithelial scatter protein.";
RL      Biochem. Biophys. Res. Commun. 168:1082-1086(1990).
RN      [3]
RP      VARIANTS EPPK LYS-160; GLN-162 AND TRP-162.
RX      MEDLINE=94214498; PubMed=7512862;
RA      Reis A., Hennies H.-C., Langbein L., Digweed M., Mischke D.,
RA      Dreschler M., Schroek E., Royer-Pokora B., Franke W.W., Sperling K.,
RA      Kuester W.;
RT      "Keratin 9 gene mutations in epidermolytic palmoplantar keratoderma
RT      (EPPK).";
RL      Nat. Genet. 6:174-179(1994).
RN      [4]
RP      VARIANTS EPPK TRP-162 AND SER-167.
RX      MEDLINE=95164963; PubMed=7532199;
RA      Rothnagel J.A., Wojcik S., Liefer K.M., Dominy A.M., Huber M.,
RA      Hohl D., Koop D.R.;
RT      "Mutations in the 1A domain of keratin 9 in patients with
RT      epidermolytic palmoplantar keratoderma.";
RL      J. Invest. Dermatol. 104:430-433(1995).
CC      -1- FUNCTION: MAY SERVE AN IMPORTANT SPECIAL FUNCTION EITHER IN THE
CC      MATURE PALMAR AND PLANTAR SKIN TISSUE OR IN THE MORPHOGENIC
CC      PROGRAM OF THE FORMATION OF THESE TISSUE.
CC      -1- SUBUNIT: HETEROTETRAMER OF TWO TYPE I AND TWO TYPE II KERATINS.
CC      -1- TISSUE SPECIFICITY: EXPRESSED IN THE TERMINALLY DIFFERENTIATED
CC      EPIDERMIS OF PALMS AND SOLES.
CC      -1- DISEASE: DEFECTS IN KRT9 ARE A CAUSE OF EPIDERMOLYTIC PALMOPLANTAR
CC      KERATODERMA (EPPK), AN AUTOSOMAL DOMINANT DISEASE CHARACTERIZED BY
CC      DIFFUSE THICKENING OF THE EPIDERMIS ON THE ENTIRE SURFACE OF
CC      PALMS AND SOLES SHARPLY BORDERED WITH ERYTHEMATOUS MARGINS.
CC      -1- MISCELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND
CC      MICROFILAMENTAR KERATIN: I (ACIDIC; 40-55 KDa) [K9 TO K20] AND II
CC      (NEUTRAL TO BASIC; 56-70 KDa) [K1 TO K8].
CC      -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC      -1- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE A 60 KDa CHAIN OF
CC      PLACENTAL SCATTER PROTEIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial

```



```

775 AspProAlaValSer..... 779
1194 TCATGCAAAAGTTAGAGATTATGCCGAACGCTTTAAAAATTAGCT 1243
780 .....GLYT 781
1244 ATGACGAGATTAATTATATAGCAAAATATACGCGGAGACGAGCACTATT 1293
781 yrAsp.....ValLeuArgValGlnGlyThrThrThrThrVal 793
1294 GATGATCGGTTACTGCAATTAATACCGCATGCGCGCT..... 1332
794 ValAlaGlnThrThrValProThrValThrLeuSerGlyLeuThrProse 810
1333 .....ATTGCTGGCGTGTGT 1348
810 rThrAlaTyrThrTyrAlaValArgAlaLysAsnValAlaGlyAspVal 827
1349 CTGCTGCTGCGCGATTTAATCAATTTGAAAAAGTTAAACATTAATCTTGT 1398
827 erAlaLeuSerAlaProValThrPhe..... 835
1399 ATCAGCAATAGCAAAAAAGAAAGTGCACCATTCAAAAGTGGTTCGAGA 1448
836 ...ThrThrAlaAlaProProValAspThrVal..... 845
1449 GCGCATTTTGGTAAGAGAGTGCCTAATTTAAAGCAACTAAAGATGAGA 1498
846 .....AlaProThrValProGlyThrProValAlaSerAsn.... 857
1499 AAATCGAAGAAATCATCGTCAAAATGCGAGCGGATCACTCAAAAGCAA 1548
858 .....ValAlaThrThrGlyAlaThrLeuThrThrThrAla 869
1549 GTTGATGATCTTATCGCAAAAGTACGCGCAAAATTAACCAAGATGAGCT 1598
870 SerThrAsp.....SerGlyGlySerGly..... 877
1599 ATCAAAAGTTGTGATACATGATGAATGCTCAAAACATAGCAAAAGTGA 1648
878 .....LeuAlaGlyTyrGluValLeuArg.....ValS 887
1649 CAACACGCTTAGATTAATGTTATCTCATCTGTAAGTCATTACC..... 1692
887 erGlyThrThrGlnThrLeuValAlaSerProThrThrAlaThrValAla 903
1693 .....TCGTC 1697
904 LeuAlaGlyLeuThrProAlaThrAlaTyrSerTyrValValaArgAla 920
1698 TTAATGATTCGAAATGTATTA.....GTGCGTCCACCTT 1732
920 sAspGlyAlaGlyAsnValSerAlaValaSerSerProValThrThrThr 937
1733 CAATGTTGATCAAAAGTTATCTTCTCTTCATTAATTTAGGAGATCTAG 1782
937 hrLeuProValThrSerThrProSerGlyThrValValTyrSerThrAsn 953
1783 CATTTGAGTACGCGCTGCGCGCTGCGAGCGGTTCTCAAGATTGGACCTA 1832
954 SerTyrAsnValGlyPhe.....ThrGlySerValLysIleThr.. 966
1833 CGGCGCTCGCGCGGTGGCTACCGACGATGGAGCTAGCGCGCTGGGCC 1882
967 .....AsnThrGlyThrThrProLeuThrThrThrLeuGlyPheAla 981
1883 CTGCGACGCGGTAGCCAGATTGAGCTACGCGCTCGCGGTGATCT 1932
981 heProSerGlyGlnGln.....ValThrGlnGlyTyrPser 992
1933 CAGCATTGAGACTAC...GGCTGCGCGCTGCGAGCGGTTCTCAAGATTG 1979

```

```

993 AlaThrTyrPserGlnThrGlyThrThrValThrAlaThrGlyLeuSerTyr 1009
1980 GAGCTACGCGCTGCGCGGTGGCTGCTACGACGATGGAGCTAGCGCC 2029
1009 pAsnAlaThrLeuGlnProGlyGlnSerThrAspIleGlyPheAsnGlyS 1026
2030 TGGCGCCCTGCGAGCGGTAGCCCAAGATTGAGCTAC 2064
1026 erHisProGlyThrAsnThrAsnProAlaSerPhe 1037
seq_name: SwissProt_39:TALI_MOUSE
seq_documentation_block:
ID TALI_MOUSE STANDARD; PRT; 2541 AA.
AC P26039;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TALIN.
GN TALN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-FTBROBLAST;
RX MEDLINE-91015390; PubMed-2120593;
RA Rees D.J.G., Ades S.A., Singer S.J., Hynes R.O.;
RT Nature 347:685-689(1990).
CC -!- FUNCTION: PROBABLY INVOLVED IN CONNECTIONS OF MAJOR CYTOSKELETAL
CC STRUCTURES TO THE PLASMA MEMBRANE. HIGH MOLECULAR WEIGHT
CC CYTOSKELETAL PROTEIN CONCENTRATED AT REGIONS OF CELL-SUBSTRATUM
CC CONTACT AND IN LYMPHOCYTES, AT CELL-CELL CONTACTS.
CC -!- SUBUNIT: BINDS WITH HIGH AFFINITY TO VINCULIN AND WITH LOW
CC AFFINITY TO INTEGRINS.
CC -!- PTM: PHOSPHORYLATED.
CC -!- SIMILARITY: CONTAINS A DOMAIN FOUND IN BAND 4.1, EZRIN, MOESIN,
CC RADIOLIN, AND TALIN.
CC -!- SIMILARITY: IN THE C-TERMINUS TO YEAST SLA2 AND C.ELEGANS ZK670.3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X56123; CAA39588.1; -
CC PIR; S11661; S11661.
CC HSSP; P04002; 1WFA.
CC MGD; MGT:1099832; TLN.
CC INTERPRO; IPR000299; -
CC INTERPRO; IPR002358; -
CC PFAM; PF00373; Band_41; 1.
CC PROSITE; PS00660; BAND_41_1; 1.
CC PROSITE; PS00661; BAND_41_2; 1.
CC PROSITE; PS00657; BAND_41_3; 1.
CC Structural protein; Cytoskeleton; Phosphorylation.
CC FT DOMAIN 173 332
CC FT VARIANT 1105 1105 L -> P.
CC FT VARIANT 2180 2180 K -> M.
CC SO SEQUENCE 2541 AA; 269832 MW; 14EF5ABE9FC2CB CRC64;
alignment_scores:
Quality: 159.00 Length: 902
Ratio: 0.408 Gaps: 39
Percent Similarity: 43.237 Percent Identity: 20.067
alignment_block:

```



```

1223 euleuSerAspLeuLeuProSerThrGlyThrPheGlnGluAlaGln 1239
1387 CATAACTTTCATCAGCAATAGCAAAAAGAGAAATGACATTCAAA 1436
1240 SerArgLeu..... 1242
1437 CTGGTTCGAGAGCGTGATTGCTTAAGAAAGTCCTAATTAATTAACAA 1486
1243 .....AsnGluAla.....AlaAlaG 1248
1487 CTAAAGATGAGAAAATGAAAGAAATCATCGCTCAAAATGCGAGCGATC 1536
1248 LysLeuasnGlnAlaAlaThrGlnLeuValGln..... 1258
1537 ACCCAAGAGCAAGTTGATGATCTAATGCAAAAGTAAGGCAAAATTAC 1586
1259 AlaserArgGlyThrProGlnAspLeuAlaArgAlaSerGlyArgPheGln 1275
1587 CCAAGAT.....GAGC 1597
1275 YGlnAspPheSerThrPheLeuGlnAlaGlyValGlnMetAlaGlyGln 1292
1598 TAACAAGAGTTGTTGATAGTATGATGCTCAACATAGCAAAATGTG 1647
1292 LaproSerGlnGluAspArgAlaGlnAlaValSerAsnLeuLysGlyLe 1308
1648 ACAAGAGCTTAGATAGTAAATCTCATCTGTAAGTCAATTACCTGCTC 1697
1309 SerMetSerSerLysLeuLeuLeuAlaAlaLysAlaLeuSerThrAs 1325
1698 TAATGATTCAGAAATGTA.....TTAGTGGCTCCAACTTCAATGT 1738
1325 ProAlaSerProAsnLeuLysSerGlnLeuAlaAlaAlaArgAlaVal 1342
1739 TGGATCAAACTTATCTCTCTT.....CAATTGCTCAGG 1773
1342 alThrAspSerIleAsnGlnLeuIleThrMetCysThrGlnAlaPro 1358
1774 GGAATCTAGATTGAGCTAGCGCGCTCGC..... 1803
1359 GlnGlnLysGlnCysAspAsnAlaLeuArgGlnLeuGlnThrValArgGln 1375
1804 .....CTGGCAGCGGTTCTCAAGATTGAGCTAC...GGCC 1837
1375 uLeuLeuGlnAsnProValGlnProIleAsnAspMetSerThrPheGly 1391
1838 TCGCTCCGGGTGCTAGCCAGC.....ATTGAGCTAGCGGCTG 1878
1392 CysLeuAspSerValMetGlnAsnSerLysValLeuGlnAlaMetThr 1408
1879 CGCCCTGCGAGCGGTAGCCAAAGATTGA..... 1906
1408 rGlyIleSerGlnAsnAlaLysasnGlnLysLeuProGlnPheGlnYAsp 1425
1907 .....GCTACGGCCCTGCGCTCGGTTGATCTCAGCAATTGAGCTAGCGCC 1951
1425 lalIleAlaThrAlaSerLysAlaLeu.....CysGlyPheThrGln 1438
1952 TGGCCCTGCGAGCGGTCTCAAGATTGA..... 1991
1439 AlaAlaAlaGlnAlaAlaIleAlaThrLeuValGlnValSerAspProAsnSerGln 1455
1982 .....GCTACGG 1988
1455 nAlaGlyGlnGlnGlyLeuValGlnProThrGlnPheAlaArgAlaAsnG 1472
1989 CTTGCGCTCGGCTGCTAGCCAGCATTTGAGCTAGCGCTGC..... 2032
1472 lmalIleGlnMetAlaCysGlnSerLeuGlnGlnProGlnCysThrGln 1488
2033 .....GCCCTGCGAGCGGTAGCCCAAGATTGAGCTAGCGC 2067
1489 AlaglnValLeuSerAlaAlaThrIleValAlaLysHisThrSerAlaLe 1505

```

2068 CTGC 2071
 111
 1505 uCys 1506

seq_name: SwissProt_39:FLIC_SALON

seq_documentation_block:

ID FLIC_SALON STANDARD; PRT: 507 AA.

AC 006974:

DT 01-OCT-1996 (rel. 34, Created)

DT 01-OCT-1996 (rel. 34, Last sequence update)

DT 30-MAY-2000 (rel. 39, Last annotation update)

DE FLAGELLIN (PHASE-1-C FLAGELLIN).

GN FLIC.

OS Salmonella oranienberg.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Salmonella.

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN-AFCC 9239;

RX MEDLINE=93374829; PubMed=7690024;

RA Masten B.J., Joys T.M.;

RT "Molecular analyses of the Salmonella g... flagellar antigen

RT complex.";

RL J. Bacteriol. 175:5359-5365(1993).

CC -1- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO

CC FORM THE FILAMENTS OF BACTERIAL FLAGELLA.

CC -1- MISCELLANEOUS: INDIVIDUAL SALMONELLA SEROTYPES USUALLY ALTERNATE

CC BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLA, TERMED

CC PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

CC EMBL: 215070; CAA78779.1; -
 DR INTERPRO: IPR001029; -
 DR INTERPRO: IPR001492; -
 DR PFAM: PF00700; Flagellin_C; 1.
 DR PFAM: PF00669; Flagellin_N; 1.
 DR PRINTS: PR00207; FLAGELLIN.
 KW Flagella.
 FT INIT_MET 0
 FT SEQUENCE 507 AA; 53065 MW; 17A88B4BBA4A6125 CRC64;

alignment_scores:

Quality: 157.50 Length: 501

Ratio: 0.592 Gaps: 23

Percent Similarity: 53.094 Percent Identity: 21.557

alignment_block:

US-09-306-689-12 x FLIC_SALON ..

Align seg 1/1 to: FLIC_SALON from: 1 to: 507

```

469 CGCAATATATATGTCACACGCTCAACAGGTTAGCAGCATTCACACCGC 518
14 GlnAsnAsnLeuAsnLysSerGlnSerLeu.....SerSerAl 27
519 TATTGCTTAACGAGCGGTGCAATGTGTATTCGCTCCACAAATTGATA 568
27 alleGluArgLeuSerSerGlyLeuArgIleAsnSerAlaLysAspAsp 43
569 AATGCTACAGAAACCTAAAGCAGGCAAGCATTA..... 603
44 .....AlaAlaGlyGlnAlaIleAlaIleAlaAsnArgPheThr 54

```

```

604 GGTCTGCCGAAAGCATTTGACAA.....ATGCAATATAA..... 639
      : : : : : : : : : : : : : : : : : : : : : : : :
55 SerAsnIleLysGlyLeuThrGlnAlaSerArgAsnAlaAsnAspIleYl 71
      : : : : : : : : : : : : : : : : : : : : : : : :
640 .....GCCAAACTGTATTATTCGGCATTCATCTATTATTAGGCTCAG 682
      : : : : : : : : : : : : : : : : : : : : : : : :
71 eSerIleAlaGlnThrThrGluGlyAlaLeuAsnGlnIleAsnAsnAsn 88
      : : : : : : : : : : : : : : : : : : : : : : : :
683 TATTGGCTGCAATGCATTAGATGAGGCTTACAGAAATAACAGCAACCA 732
      : : : : : : : : : : : : : : : : : : : : : : : :
88 euGlnArgValArgGluLeuSerValGlnAlaThrAsnGlyThrAsnSer 104
      : : : : : : : : : : : : : : : : : : : : : : : :
733 CATGCTCTGCTAAAGCTGCTGGAGCTAACAAATTCATTATTATTAATA 782
      : : : : : : : : : : : : : : : : : : : : : : : :
105 AspSerAspLeuLysSerIleGlnAspGluIleGlnGlnArgLeuGlu 121
      : : : : : : : : : : : : : : : : : : : : : : : :
783 TATT.....GCTAATTCGATAAAACACTTG 808
      : : : : : : : : : : : : : : : : : : : : : : : :
121 uIleAspArgValSerAsnGlnThrGlnInpHeAsnGlyValLysValLeu 138
      : : : : : : : : : : : : : : : : : : : : : : : :
809 AGCAATTTGTCAGCAAAATTAGTCAATTGTTCAAA..... 846
      : : : : : : : : : : : : : : : : : : : : : : : :
138 eArgLAspAsnGlnMetLysIleGlnValGlyAlaAsnAspGlyGluThr 154
      : : : : : : : : : : : : : : : : : : : : : : : :
847 .....CTACAA.....AATATCAAGGCTTAGG..... 870
      : : : : : : : : : : : : : : : : : : : : : : : :
155 ILeThrIleAspLeuGlnLysIleAspValLysSerLeuGlyLeuAspI 171
      : : : : : : : : : : : : : : : : : : : : : : : :
871 .....ACTTTAGAGAC.....A 883
      : : : : : : : : : : : : : : : : : : : : : : : :
171 yPheAsnValAsnGlyProLysGlnAlaThrValGlyAspLeuLysSer 188
      : : : : : : : : : : : : : : : : : : : : : : : :
884 AACTCAAAATATCGTGACATTGATAAGCGCTTGCT..... 924
      : : : : : : : : : : : : : : : : : : : : : : : :
188 ePheLysAsnValThrGlyTyArgPThrTyArgValGlyAlaAsnLys 204
      : : : : : : : : : : : : : : : : : : : : : : : :
925 .....TTAGATTTATCTCAGGCGCTATTACGGCGCAACAGCTGCAT 968
      : : : : : : : : : : : : : : : : : : : : : : : :
205 TyArgValAspValAsnSerIleAlaValIleThrAspThrThrAla 221
      : : : : : : : : : : : : : : : : : : : : : : : :
969 TGTACTTGCAGATAA.....AATGCTTCAACAGCTAAATAAGTGC 1009
      : : : : : : : : : : : : : : : : : : : : : : : :
221 GThrValProAspLysValTyValAsnAlaAlaAsnGlyGlnLeuThr 238
      : : : : : : : : : : : : : : : : : : : : : : : :
1010 GTGGCGGTTTGAATTGGCAACCAAGTGTGTGTAATTTACCAAGCC 1059
      : : : : : : : : : : : : : : : : : : : : : : : :
238 hrAla.....AspAlaGlnAsnAsnThrAlaValAspLeuPheLysSer 252
      : : : : : : : : : : : : : : : : : : : : : : : :
1060 GTTCTCTTACATTTTAGCCCAAGCTGTTGCAGCAGCTTATCTTCAC 1109
      : : : : : : : : : : : : : : : : : : : : : : : :
253 ThrLysSer.....AlaAlaGlyThrAspAspAl 262
      : : : : : : : : : : : : : : : : : : : : : : : :
1110 TGGGCGCTGCGCTGCTTAATTGCTTCTACTGTTCTCTGCGATTAGCC 1159
      : : : : : : : : : : : : : : : : : : : : : : : :
262 aLysAlaIleAlaThrSerIleLysGly..... 271
      : : : : : : : : : : : : : : : : : : : : : : : :
1160 CATTTAGCATTTGCCGATTCGCCGATAAATTAATCATCAAAAGTTTA 1209
      : : : : : : : : : : : : : : : : : : : : : : : :
272 .....GlyLysValGlyAspThrPheAsp..... 279
      : : : : : : : : : : : : : : : : : : : : : : : :
1210 GAGAGTTATGCCGACGCTTAAAAAATTAGGCTATGACGAGATAATT 1259
      : : : : : : : : : : : : : : : : : : : : : : : :
280 .....TyTrpSerGlyValSerPheThrIleAspThr 290
      : : : : : : : : : : : : : : : : : : : : : : : :
1260 ATTACAGAAATATCAGCGGGAACAGGACTATTATGATCGATCGTTACTG 1309
      : : : : : : : : : : : : : : : : : : : : : : : :
290 sAlaGlySer.....AspGlyAsnGlyThrValSerThrThrIleAsnG 305
      : : : : : : : : : : : : : : : : : : : : : : : :
1310 CAATTAAATACCCATTGGCCGCTATTGCTGGTGCTGCTCTGCTGCA 1359
      : : : : : : : : : : : : : : : : : : : : : : : :
305 yGluLysValThrLeuThr.....IleSerAspIleGlyAlaSerAla 319

```

```

1360 GCCGATTTAACATTGAAAAAGTTAAACATATCTTGTCATCGAATAG 1409
      : : : : : : : : : : : : : : : : : : : : : : : :
320 ThrAspValAsnSerAlaLysIleGln.....Se 329
      : : : : : : : : : : : : : : : : : : : : : : : :
1410 CAAAAAGAGAAAGTGACC.....ATTCAAAACGTGTTCCGAGAGCTG 1453
      : : : : : : : : : : : : : : : : : : : : : : : :
329 rSerLysAspValTyThrSerValAlaSerGlyGlnPheThrPheAla 346
      : : : : : : : : : : : : : : : : : : : : : : : :
1454 ATTTGCTTAA.....GAAGTGCCTAATTATTAAGCACTATAA 1491
      : : : : : : : : : : : : : : : : : : : : : : : :
346 sPlyThrLysAsnGlnSerAlaLysLeuSerAspLeuGluAlaAsn 362
      : : : : : : : : : : : : : : : : : : : : : : : :
1492 GATGAGAAAATGCAAGAAATTCGCTCAAAATGCGGAGCGATCCAC 1541
      : : : : : : : : : : : : : : : : : : : : : : : :
363 AlAlaValLysGlyGluSerLysIleThrValAsnGlyAlaGluThrAl 379
      : : : : : : : : : : : : : : : : : : : : : : : :
1542 AAAGCAAGTGAATGATCTTATCGCAAAAGTAAAGGCAAAATTCAC 1591
      : : : : : : : : : : : : : : : : : : : : : : : :
379 aAsnAlaAlaGlyAspLysValThrLeuAlaGlyLysThrMetPheIle 396
      : : : : : : : : : : : : : : : : : : : : : : : :
1592 ATGAGCTATCAAAAGTTGT.....GATAACTATGAA 1623
      : : : : : : : : : : : : : : : : : : : : : : : :
396 sPlyThrAlaSerGlyValSerThrLeuIleAsnGluAspAlaAla 412
      : : : : : : : : : : : : : : : : : : : : : : : :
1624 TTGCTCAACATAGCAAAATGTGCACAAACAGCTTATGATATATCTC 1673
      : : : : : : : : : : : : : : : : : : : : : : : :
413 AlAlaLysSerThrAlaAsnProLeuAlaSerIleAspSerAlaLeu 429
      : : : : : : : : : : : : : : : : : : : : : : : :
1674 ANCTGTAGTGCATTTACCTCGCTAATGATTCG...AGAAATGTATTA 1720
      : : : : : : : : : : : : : : : : : : : : : : : :
429 rLysValAspAlaValArgSerSerLeuGlyAlaIleGlnAsnArgPhe 446
      : : : : : : : : : : : : : : : : : : : : : : : :
1721 TGGCTCAACTTCATGTTGATTCACAAAGTTTATCTCTTCATTTGCT 1770
      : : : : : : : : : : : : : : : : : : : : : : : :
446 sPserAlaIleThrAsnLeuGlyAsnThrValThrAsnLeuAsnSerAla 462
      : : : : : : : : : : : : : : : : : : : : : : : :
1771 AGG 1773
463 Arg 463

seq_name: SwissProt_39:Y109_YEAST
seq_documentation_block:
ID Y109_YEAST STANDARD: PRT: 995 AA.
AC P40442.
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOPHETICAL 99.7 KDA PROTEIN IN SDL1 5'REGION PRECURSOR.
GN Y11169C OR Y19402.07C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RA Barrett B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jørgels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.,
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: Z46921; CAA87023.1; -
DR

```



```

422 TAGTCAAAAGCGCCGAGAGTTGGGATTGAG.....GTACAAAGA 462
      :::::::::: :: :::::::::: ::::::::::
900 laThraenAlaAlaAlaGlnAsnAlaIleLysLysLysLeuValGlnArg 916
      ::::::::::
463 GAAGAACGC.....AATATATTGC 482
      ::::::::::
917 LeuGlnHisAlaAlaLysGlnAlaAlaSerAlaThrGlnThrIleAl 933
      ::::::::::
483 AACAGCTCAACACGTTTAGCAGCATTCAAACGCTATTGGC..... 525
      ::::::::::
933 aAlaIleGlnHisAlaAlaSerThrProLysAlaSerAlaGlyProGlnP 950
      ::::::::::
526 ..TTAACTGACCGCTGGCATTTGTATTACCGTCCACAATTTGATTAATG 573
      ::::::::::
950 roLeuLeuValGlnSerCysLysAlaValAlaGlnGlnIleProLeuLeu 966
      ::::::::::
574 CTACAGAAACTAAAGCAGCGCCAGCATTTAGGTTCTGCCGAAACATTGT 623
      ::::::::::
967 ValGlnGlyValArgGlySerGlnAlaGlnProAspSerProSer.... 981
      ::::::::::
624 ACAAAATGCAAAATPAAGCAAAACTGATTTATCTGGCATTCATCTATT 673
      ::::::::::
982 .....AlaGlnLeuAlaLeuIleAlaAlaSerGlnSerPhe. 993
      ::::::::::
674 TAGGCTCAGATTGGCTGGAATGGATTAGATGAGCGCTTACAGAAATAC 723
      ::::::::::
994 .....LeuGlnProGly 997
      ::::::::::
724 AGCACCACATGCTCTTGTCTAAAGCTGGCTTGAGCTAACAAATTCATT 773
      ::::::::::
998 GLyLysMetValAlaAlaAlaLysAla..... 1006
      ::::::::::
774 AATGAAATATTCTCTAATTCAGTAAACACCTGACGAATTTGGTGAG. 822
      ::::::::::
1007 .....SerValProThrIleGlnAspGlnAlaSerAl 1017
      ::::::::::
823 .....CAAATAGTCATTTGGTCAAACTACAAAT...ATCMAAGC 864
      ::::::::::
1017 laMetGlnLeuSerGlnCysAlaLysAsnLeuGlnThrAlaLeuAlaGln 1033
      ::::::::::
865 TTAGAGCACTTAGAGACAAACCTCAAAATATGCGTGGACTGTAAAGC 914
      ::::::::::
1034 LeuArgThrAlaAlaGlnLysAlaGlnLysGlyProLeuGlnLys 1050
      ::::::::::
915 TGCGCTTGATTAGATGTATCTCAGGCGTA.....TTATCGGCGC 955
      ::::::::::
1050 tAspSerAlaLeuSerValValGlnAsnLeuGlnLysAspLeuGlnGluV 1067
      ::::::::::
956 CAACAGCTGCACCTTGACTTGACAGATAAAATGCTTCAACAGCTAAAAA 1005
      ::::::::::
1067 alLysAlaAlaAla.....ArgAspGlyLysLeuLysProLeu 1079
      ::::::::::
1006 GTGGTGCGGGTTTGAATTGGCAACCAAGTTTGTGTAATATTACCAA 1055
      ::::::::::
1080 proGlyGlnThrMetGlnLysCysThrGlnAspLeuGlnLysSerThrLy 1096
      ::::::::::
1056 AGCGCTTCTTCTTAC..... 1071
      ::::::::::
1096 sAlaValSerSerAlaIleAlaGlnLeuLeuGlyValAlaAlaGlnLys 1113
      ::::::::::
1072 .....ATTTAGCCCAACGCTGTTCAGCAGAGTTTATCT 1104
      ::::::::::
1113 snGlnAsnThrAlaGlyIleAlaAlaArgAspValAlaGlyLysLeuArg 1129
      ::::::::::
1105 TCAACTGGGCTGGCTGCTTAAATGCTTCT...ACTGTTTCTCTTGC 1151
      ::::::::::
1130 SerLeuAlaGlnAlaAlaArgGlyValAlaAlaLeuThrSerAspProAl 1146
      ::::::::::
1152 GATTAGCCCAATTAGCATTTGCCGATTTGCCGATAAATTAATCATGCAA 1201
      ::::::::::
1146 aValGlnAlaIleValLeuAspThrAlaSerAspValLeuAspLysAlaAs 1163
      ::::::::::
1202 AAAGTTTAGAGATTATGCGAACGCTTTAAAAAATTAGGCTATGACGGA 1251
      ::::::::::
      ::::::::::
1163 eSerLeuIleGlnGluAlaLys.....LysAlaAlaGlnHisProGly 1177
      ::::::::::
1252 GATAATTATTATGACGAATATACAGCGGGGACAGAGCACTATGATGATC 1301
      ::::::::::
1178 Asp.....ProGlnSerGlnGlnArgLeuAlaGlnValAlaLysAl 1191
      ::::::::::
1302 GGTTACTGCA..... 1311
      ::::::::::
1191 aValThrGlnAlaLeuAsnArgCysValSerCysLeuProGlyGlnArg 1208
      ::::::::::
1312 ..ATTAAATACCGCATTTGGCGCTATTGCTGGTGCTGTCTGTCTGCA 1359
      ::::::::::
1208 spValAspAsnAlaLeuArgAlaValGlyAspAlaSerLysArgLeuLeu 1224
      ::::::::::
1360 GCCGATTTA.....ACATTGAAAAAGTTAAACATA 1391
      ::::::::::
1225 SerAspLeuLeuProProSerThrGlyThrPheGlnGlnAlaGlnSerAr 1241
      ::::::::::
1392 TCTTGTCATCAGCAATAGCAAAAAAGAGAAAGTGACCATTCAAAACGTGT 1441
      ::::::::::
1241 gLeu..... 1242
      ::::::::::
1442 TCCGAGAGCGCTGATTTTGTCTAAAGACTGCCATTAATTAAAGCACTAAA 1491
      ::::::::::
1243 ..AsnGluAla.....AlaAlaGlyLeu 1249
      ::::::::::
1492 GATGAGAAATCGAAGAAATCATCGCTCAAAATGGCGAGCGGATCACCTC 1541
      ::::::::::
1250 AsnGlnAlaAlaThrGlnLeuValGln.....AlaSe 1260
      ::::::::::
1542 AAACCAAGTTGATGATCTTATCGCAAAAGATACGCCAAATATACCAAG 1591
      ::::::::::
1260 rArgGlyThrProGlnAspLeuAlaArgAlaSerGlyArgPheGlyLysAl 1277
      ::::::::::
1592 AT.....GAGCTATCA 1602
      ::::::::::
1277 spHeSerThrPheLeuGlnAlaGlyValGlnMetAlaGlnAlaPro 1293
      ::::::::::
1603 AAAGTTGTGATACCTATGAAATTCCTCAACATAGCAAAATATGACAAA 1652
      ::::::::::
1294 SerGlnLysAspArgAlaGlnValValSerAsnLeuLysGlyIleSerMe 1310
      ::::::::::
1653 CAGCTGATATAGTAAATCTCATCTCAATGCTAAGCTTATACCTCGCTATG 1702
      ::::::::::
1310 tSerSerLysLeuLeuLeuAlaAlaLysAlaLeuSerThrAspProA 1327
      ::::::::::
1703 ATTCGAGAAATGTA.....TTAGTGCCTCCAACTTCAATGTTGAT 1743
      ::::::::::
1327 laAlaProAsnLeuLysSerGlnLeuAlaAlaAlaArgAlaValThr 1343
      ::::::::::
1744 CAAGTTTATCTTCTCTT.....CAATTGCTAGGGATC 1778
      ::::::::::
1344 AspSerIleAsnGlnLeuIleThrMetCysThrGlnGlnAlaProGlyG 1360
      ::::::::::
1779 TCACGATTGGAGCTACGCGCTGCGC..... 1803
      ::::::::::
1360 nLysGlnCysAspAsnAlaLeuArgGlnLeuGlnThrValArgGlnLeuL 1377
      ::::::::::
1804 .....CCTGCGAGCGGTTCTCAAGATTGGAGCTAC...GGCTGCGT 1842
      ::::::::::
1377 euGlnAsnProValGlnProIleAsnAspMetSerLysPheGlyCysLe 1393
      ::::::::::
1843 CCGGCTGGCTCTAGCCAGC.....ATTGAGCTACGCGCTGCGGCC 1883
      ::::::::::
1393 uAspSerValMetGlnAsnSerLysValLeuGlyGlnAlaMetThrGly 1410
      ::::::::::
1884 TGGCAGCGGTAGCCAGATTGGA..... 1906
      ::::::::::
1410 leSerGlnAsnAlaLysAsnGlyAsnLeuProGlnPheGlyAspAlaIle 1426
      ::::::::::
1907 GCTACGGCCTGCGCTCGCGTGAGATCTCAGCATTTGACCTACGCGCTGCGC 1956
      ::::::::::
      ::::::::::

```


alignment_block:

US-09-306-689-12 x FLA3_CAMJE ..

Align seg 1/1 to: FLA3_CAMJE from: 1 to: 574

```

137 GTAGCCAGATTGGAGCTACGCGCTCGCTCCGGGTGATCTC..... 178
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
      ValAlaAlaLeuAsnAlaLysAla.....AsnAlaAspLeuAsnSerLy 22
179 .ACATTGGAGCTACGCGCCCGCCGCGAGCGGTTCCAGATGAGC 227
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
22 sSerLeuAspAlaSerLeuSerArgLeuSerSerGlyLeuArgIleAsn 39
228 CTACGCGCTCGCTCGCGGTGCTCTAGCC.....ACGATTGGAGCT 268
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
39 erAlaAlaAspAlaSerLySerGlyMetAlaIleLysAspSerLeuArgSer 55
269 ACGGCGCTGCGCCCTGGCAGCGGTAGCCAGATTGGAGCTACGCGCTGCGT 318
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
56 GluAlaAsnThrLeuGlyGlnAlaIleSerAsnGlyAsnAspAlaLeuG 72
319 CCGGCTGAGCTAGCTTCCAAAAACTGGGGCAA.....AAAAATTAT 362
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
72 yIleLeuGlnThrAlaAspLysAlaMetAspGluGlnLeuLysIleLeuA 89
363 CCTCTATATTCGCCAAATACCAATATGATGACAGCAAGTATGCTT 412
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
89 sphrIleLysThrLysAlaThrAlaAla..AlaGlnAspGlyGlnSerL 105
413 TACAGGATTTAGTCAAAGCGCGCAGAGATTGGGATTGAGTCAAGA 462
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
105 eu.....LysThrArgThrMetLeuGlnAlaAspIleAsnArg 117
463 .....GAAGACGCAATATATTTGCA..... 483
118 LeuMetGlnIleuLeuAspAsnIleAlaAsnThrThrSerPheAsnGly 134
483 ..... 483
134 sGlnLeuLeuSerGlyAsnPheIleAsnGlnGlnPheGlnIleGlyAla 151
484 .....ACAGCTCAAAACAGTTTAGCAGCAGATTCAACCGCT... 519
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
151 erSerAsnGlnThrValLysAlaSerIleGlyAlaThrGlnSerSerLys 167
520 ATGGCGCTTAAC.....GAGCGTGCAATGTGTATCCGCTCA..... 558
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
168 IleGlyLeuThrArgPheGlnThrGlySerArgIleSerValGlyGly 184
559 ..... 559
184 uValGlnPheThrLeuLysAsnTyraSnglyIleAspAspPheLysPheG 201
566 ATAAATGCTACAGAAAACCTAAAGCAGCGCAAGCATTAGCTTCT...GCC 612
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
201 InLysValAlaIleSerThrSerValGlyThrGlyLeuGlyAlaLeuAla 217
613 GAAGAGCTTGTACAAATGCAAAATAAAGCCAAACTGTATTATCTGGCAT 662
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
218 AspGlnIleAsnLysAsnAlaAspLys.....ThrGlyVala 229
663 TCATCTATT.....TTAGGCTCAGTATTGGCTG 691
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
229 IArgAlaIleThrPheThrValGlnThrArgGlyMetGlyAlaValArgLag 246
692 GAATGATTTAGATGAGCGCTTACAGAAATAACAGCAACCAATGCTCTT 741
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
246 LysAlaThrSerAspAspPheAlaIleAsn..... 255
742 GCTAAAGCTGCGCTTGAGCTA.....AC 764
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
256 .....GlyValLysIleGlyLysValAspTyryLysAspGlyAspAl 269

```

```

765 AAATTCATTATTTGAAAAATATTCATTATTCAGTAAAA..... 801
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
269 aAsnGlyAlaLeuValSerAlaIleAsnSerValLysAspThrThrGly 286
802 .....ACACTTGACGAATTGGTGAGCAAAAT.....AGTCATTT 837
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
286 aLglAlaSerIleAspGluAsnGlyLysLeuLeuThrSerArgGlu 302
838 GGTCAAAACTACAAATATCAAAAGCCTTAGGCACTTATGAA..... 879
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
303 Gly.....ArgGlyIleLysIleGlyLysAsnIleGlyArgGlyAl 316
880 .....GACAACTCAAAAATATCGGTGACCTTGATTA... 912
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
316 aPheIleAsnProAsnMetLeuGlnAsnTyryGlyArgLeuSerVal 333
913 .....GCTGGCCTTGCTTATGATGATTATC 936
333 yAsnAspGlyLysAspIleLeuIleSerGlyThrAsnLeuSerAlaIle 349
937 .....TCAGGCTATTTATCGGCGCAACAGCTCAGCTTGACT 974
350 GlyPheGlyThrGlyAsnMetIleSerGlnAlaSerValSerLeuArg 366
975 TGCA.....GATAAAAATGCTTCAACAGCTAAAAAGTGGTG 1012
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
366 uSerLysGlyGlnIleAspAlaAsnValAlaAspAla.....M 379
1013 CGGTTTGAATGCGCAAAACCACTTGTGTATATTACCAAAAGCGCT 1062
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
379 etGlyPheAsnSerAlaAsnLys.....GlyAsnIleLeuGlyGlyTyr 393
1063 TCCTCTTACATTTTAGCCCAACGCTTGACAGCAGGTTATCTTCACTGG 1112
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
394 SerSer.....ValSerAlaIleThrMetSerSerThrG 404
1113 G..... 1113
404 ySerGlyPheSerSerGlySerGlyPheSerValGlySerGlyLysAsn 421
1114 .....CCTGGCTGCTGTTAATGCTCTACTGTT 1143
421 ySerThrGlyPheAlaAsnThrIleAlaIleSerAlaAlaSerGlnLeu 437
1144 TCCTGTCGATTAGCCCATTAAGCTTGGCGGTATGCCGATAATTTAA 1193
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
438 SerAlaValTyraSnaValSerAlaGlySerGly.....Phe 450
1194 TCATGCAAAAAGTTTAGAGATTATGCCGAACGCTTAAAAAATTAGCT 1243
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
450 rSerGlySerAsnLeuSerGlnPheAlaThrMetLysThrSerAlaGly 466
1244 ATGACGGAGATTAATTTATTGACA.....GAATTCAGCGCGGGAACAGG 1287
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
467 .....AsnThrLeuGlyValLysAspGlnThrArgLysValThr 479
1288 ACTATGATGCATGCTTACTGCAATTAATACCGCATTCGCGCTATTGCG 1337
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
480 ThrLeuLysGlyAlaMetAlaValMetAspIleAlaGlnThrAlaIle 496
1338 TGTGTGTGTGTGCTGCTGCGACCGGATTTAACATTTGAAAAAACTTAAC 1367
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
496 r...AsnLeuAspGlnIleArgAlaAspIleGlySerValGlnAsnGln 512
1388 ATATCTTGTCATCAGAAATAGCAAAAAAGAAAGAGACATF..... 1431
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
512 aThrThrThrThrIleAsnAsnIleThrValThrGlnValAsnValLysAla 528
1432 .....CAAAAGCTGTCGAGAGGCTGATTTTGTCTAAAGAGTCCATAATTA 1478
      ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
529 AlaGlnSerGlnIleArgAspValAlaAspPheAlaAlaGlnSerAlaAsn 545
1479 TAAAGCAACTAAAGATGAGAAAATCGAAGAAATCATCGGTCAAAATGCGC 1527

```

l::: :
545 rserLysala.....AsnIleuAlaGlnSerCly 555

THIS PAGE BLANK (USPTO)

OM of: US-09-306-689-12 to: PIR_66:* out_format : pfs
 Date: Mar 2, 2001 10:36 AM

About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frimel_nrp.model -DEV=rlp
 -Q=cgnt1_uspto_spool/US09306689/runatc_02032001_102825_9560/app_query.fasta_1.2389
 -DB=PIR_66 -PEM=fastan -SPEF=PIR -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -OGAPOP=4.500
 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blomsun62 -TRANS=human40.cdi
 -LIST=45 -DOCALIN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US09306689 -CGNT1_1_370 -NCPU=6 -ICPU=3
 -LONGLOG -NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-306-689-12
 Query length: 2088
 Database: PIR_66:*
 Database sequences: 195891
 Database length: 67900655
 Search time (sec): 196.850000

score list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
PIR1:OYBRC	+	212.00	199.47	0.0002	1706	Documentation
PIR2:S51672	+	206.50	194.17	0.0004	1705	cyclolysin - Bordetella pertussis
PIR2:A57036	+	182.50	168.14	0.0080	2491	adenylate cyclase hemolysin -
PIR1:KRNU2	+	172.00	168.40	0.0301	643	keratin - slime mold (Dictyostel
PIR2:CG3339	+	169.00	149.37	0.0394	5627	hypothetical protein P2462 [i
PIR2:A75564	+	169.00	159.19	0.0430	1467	conserved hypothetical protein
PIR2:A62615	+	168.50	160.31	0.0459	1190	surface protein Xf1981 [import
PIR2:T14961	+	166.50	156.48	0.0585	1525	hypothetical protein Y1044 - Y
PIR2:T153169	+	166.00	162.67	0.0632	638	cytokeratin 2 - human
PIR2:A44861	+	166.00	162.59	0.0632	645	keratin, 67K type II epidermal
PIR2:E75590	+	166.00	161.37	0.0630	756	methyl-accepting chemotaxis-rel
PIR2:A56143	+	165.00	157.48	0.0708	1109	surface-array protein homolog
PIR2:141061	+	164.00	161.68	0.0811	565	flagellin - Escherichia coli (S
PIR2:T13784	+	163.00	159.98	0.0916	622	keratin 9, type I, cytoskeletal
PIR2:T13634	+	163.00	153.19	0.0901	1510	probable minor tail protein gr
PIR2:A39199	+	161.50	154.56	0.1092	1045	endoglucanase B (EC 3.2.1.-)
PIR2:T38274	+	161.00	154.49	0.1163	990	probable transcription initiat
PIR2:A37284	+	160.50	154.80	0.1239	893	surface-array protein - Campylo
PIR2:DS2671	+	160.00	147.92	0.1299	2059	surface protein Xf1529 [import
PIR2:T13329	+	159.00	149.30	0.1478	1517	hypothetical protein P14625 [i
PIR2:S11661	+	159.00	145.35	0.1464	2541	tailin - mouse
PIR2:140711	+	158.50	152.51	0.1585	936	sapB protein - Campylobacter fe
PIR2:T34852	+	158.00	150.41	0.1680	1156	probable secreted protein - St
PIR2:S33192	+	157.50	156.24	0.1813	507	phase-1 flagellin - Salmonella
PIR2:A53465	+	157.50	156.23	0.1813	508	phase-1 flagellin - Salmonella
PIR2:FB3068	+	157.50	145.17	0.1767	2154	hypothetical protein P14625 [i
PIR2:SS0358	+	155.50	149.15	0.2294	995	hypothetical protein Y1169c -
PIR2:EB1965	+	155.50	148.61	0.2291	1068	probable outer membrane peptid
PIR2:HB1020	+	155.50	148.51	0.2291	1082	serotype-1-specific antigen, F
PIR2:JC1419	+	154.50	152.23	0.2630	587	FC gamma (19c) receptor II pred
PIR2:A47283	+	154.50	149.19	0.2602	873	calpactin - fruit fly (Drosophi
PIR2:T03099	+	154.00	127.86	0.2637	13288	mucin, submaxillary - pig

seq_name: PIR1:OYBRC

seq_documentation_block:

cyclolysin - Bordetella pertussis
 N:Alternate names: adenylate cyclase precursor; calmodulin-sensitive adenylate cyclase-
 N:Contents: adenylate cyclase (EC 4.6.1.1), calmodulin-sensitive; hemolysin
 C:Species: Bordetella pertussis
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999

C:Accession: S00893; S14100; S02389

R:Glaser, P.; Ladant, D.; Seeger, O.; Pichot, F.; Ullmann, A.; Danchin, A.
 Mol. Microbiol. 2, 19-30, 1988

A:Title: The calmodulin-sensitive adenylate cyclase of Bordetella pertussis: cloning
 A:Reference number: S00893; MUID:88216178

A:Accession: S00893

A:Molecule type: DNA

A:Residues: 1-1706 <GLA>
 A:Cross-references: EMBL:Y00545; NID:9396665; PIDN:CAAG6613.1; PID:9396666

R:Glaser, P.; Sakamoto, H.; Bellalou, J.; Ullmann, A.; Danchin, A.
 EMBO J. 7, 3997-4004, 1988

A:Title: Sequence of cyclolysin, the calmodulin-sensitive adenylate cyclase--haemoly
 A:Reference number: S02386; MUID:89091151

A:Contents: annotation; identification of adenylate cyclase--hemolysin difunctional p
 R:Mueller, H.; Gilles, A.M.; Glaser, P.; Krin, E.; Danchin, A.; Sarfati, R.; Barzu, O.
 Eur. J. Biochem. 196, 469-474, 1991

A:Title: Isolation and characterization of catalytic and calmodulin-binding domains o
 A:Reference number: S14100; MUID:91177021

A:Accession: S14100

A:Molecule type: protein
 A:Residues: 1-78, 'W', 80, 'W', 82-97, 'W', 99-139, 'W', 141-178, 'W', 180-399 <MUN>

R:Hackett, M.; Guo, L.; Shabanowitz, J.; Hunt, D.F.; Hewlett, E.L.
 Science 266, 433-435, 1994

A:Title: Internal lysine palmitoylation in adenylate cyclase toxin from Bordetella pe
 A:Reference number: A55167; MUID:95025937

A:Contents: annotation; lysine palmitoylation
 C:Comment: B. pertussis, the etiological agent of whooping cough, disrupts mammalian
 action by host cell calmodulin of the adenylate cyclase activity of bacterial cycloly
 C:Genetics: Adenylate cyclase activity is activated upon binding of calmodulin in the

A:Gene: cyaA, cya

C:Superfamily: cyclolysin; calmodulin-sensitive adenylate cyclase catalytic domain ho
 C:Keywords: calcium binding; calmodulin binding; cAMP biosynthesis; carbon-oxygen lya

F:15-328/Domain: calmodulin-sensitive adenylate cyclase catalytic domain homology <AD
 F:25-35/Region: calmodulin binding #status predicted

F:59-66/Region: nucleotide binding #status predicted
 F:544-1085/Domain: nucleolysin A homology <HLY>

F:1033-1041,1042-1050,1174-1182,1289-1297,1298-1308,1316-1324,1430-1438,1556-1564/Reg
 F:983/Binding site: palmitate (lys) (covalent) #status experimental

alignment_scores:

Quality:	Ratio:	Length:
212.00	0.763	512
Percent Similarity: 54.297	Percent Identity: 22.266	Gaps: 21

alignment_block:

US-09-306-689-12 x OYBRC ..
 Align seg 1/1 to: OYBRC from: 1 to: 1706

265 AGCTACGGCGCTGGCGCCGTCGACGCGGTACCAAGATTGGAGTCGCGCT 314

349 AATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 364

315 GCGTCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 364

364 AATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 381

365 TCGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414

381 TATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 389

415 CAGGATTTAGTCAAAAGCGCGCGGAGGATGGGATTTAGGATCAAGAGA 464

390 ATG.....ArgProSerLeuGly...AlaValGluArgG 400

465 AGAAGC.....AATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 499

400 nspSerGlyTyrAspSerLeuAspGlyValGlySerArgSerPheSerL 417

500 TGGGACGATTCACAAACCGCTATTTGCTTAACGAGCGTGGCATTTGCTTA 549


```

450 TGAGGTACAAAGAAAGCAAGCAATATATGCAACAGCTCAACAGCTT 499
      |||||
416 ..... SerL 417
500 TAGGACAGATTCAACCCGTAATGGCTTAACAGAGCGTGGCATTTGTTA 549
      |||||
417 euGIyLlValSerAspmet..... AlaIaVal 426
550 TCCGCTCCCAAAATGATTAATGCTACGAAACAAACAGAGCCCAAC 599
      |||||
427 GluaIaIaIaGluLeuGlnMetThrArgGlnValLeuHisIaGlyIaIa 443
      |||||
600 ATTAGATTCTGCCGAAGCATTTGACAAATATGCAAAATTA..... GCCA 643
      |||||
443 gGlnspspAlaGluProGlyValSerGlyAlaSerAlaHisTrpGlyG 460
      |||||
644 AAACGTATATATGCGCATTCATATTTAGGCTACAGTATGGCTGGA 693
      |||||
460 InArgAlaLeuGlnGlyAlaGlnAlaIaIaIa..... 472
694 ATGATTTAGATGAGCGCTTACAGATTAACAGCAACCAACATGCTTGC 743
      |||||
473 ..... GlnArgLeuValHisAlaIaIaLea 480
744 T..... AAGCTGGCTTGAGGCTAACAAATTCATTA 775
      |||||
480 aleuMetThrGlnPheGlyArgAlaGlySerThrAsnThrProGlnGln 497
776 TTGAAATATATTCCTAATTCAGTAAACAACTTGAGCAATTTGGTGAGCA 825
      |||||
497 laIaSerleuSerAlaIaIaValPheGlyLeuGlyGlnValaSerAla 513
826 ATTACTCAA..... TTTGCTCAAAACACAAATATTCAA 860
      |||||
514 ValaIaGlnThrValSerGlyPhePheArgGlySerSerArgTrpAla 530
861 AGGCTTAGGACTTAGAGACAACTCAAAATATCGGTGAGCTGATA 910
      |||||
530 yGlyPheGlyValaIaGlyAlaMetAlaLeuGlyGlyIaIaGly 547
911 AAGCTGGCGCTTTAGATTTATCTCAGGCGCTATTAACGCGCAGCA 960
      |||||
547 laValGlyAlaGlyMetSer..... 553
961 GCTGCACTTTGACTTCAGATTAATGCTTCAACAGCTAAACAGTGG 1010
      |||||
554 ..... leuThrAsp... AspAlaProAlaGlyGlnValaIa 565
1011 TCGGGGTTTGAATGGCAACCAAGTTGTGTATATTAACCAACGCG 1060
      |||||
565 alaIaGlyAlaGlnIleAlaLeuGlnLeuThrGlyGlnValGlnLeu 582
1061 TTTCTTTTACATTTTA..... GCCCAACGTTGACAGAGCT 1098
      |||||
582 laSerSerIleAlaLeuAlaIaIaAlaIaArgGlyValThrSerGly 598
1099 TTATCTTCACTGGCGCTGCTGCTTAAATGCTTCTACTGTTCTCT 1148
      |||||
599 leuGlnValaIaGlyAlaSerAlaGlyAlaIaIaIaGlyAlaLeuAla 615
1149 TCGGATTTAGCCATTCAGATTTGGCGGTTTGGCGTAATTAATTAACAG 1198
      |||||
615 alaIaLeuSerProMetGlnIleArgLeuValGlnGlnSerHisTyr 632
1199 CAAAAGTTAGAGATTATGCCGAACGCTTAAATAATTAAGCTATGAC 1248
      |||||
632 laAspGlnLeuAspLysLeuAlaGlnGlnSerSerAlaTyrGlyTyrGlu 648
1249 GGAGATTAATTTATGACAGATATACGCGGGAACAGGACTATTTGATGC 1298
      |||||
649 GlysAspAlaLeuAlaGlnIleuTyrArgAspLysThrAlaIaIaGln 665
1299 ATCGGTTACTGCAATTAATACCGCATTTGGCGCTATTGCTGAGTCTCT 1348

```

```

665 yalValaIaIaGlyValSerAlaValLeuSerThrValGlyAlaIaValS 682
      |||||
1349 CTGCTGCTGCAGCC 1362
      |||||
682 erIleAlaIaIaIa 686
seq_name: p1r2:A57036
seq_documentation_block:
  talin - slime mold (Dictyostelium discoideum)
  C:Species: Dictyostelium discoideum
  C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 29-Oct-1999
  C:Accession: A57036
  R:Reitmeier, M.; Gerisch, G.; Heizer, C.; Mueller-Taubenberger, A.
  J. Cell Biol. 129, 179-188, 1995
  A:Title: A talin homologue of Dictyostelium rapidly assembles at the leading edge of
  A:Reference number: A57036; MUID:95213284
  A:Accession: A57036
  A:Status: Preliminary
  A:Molecule type: DNA
  A:Residues: 1-2491 <RRE>
  A:Cross-references: GB:U14576; NID:g797298; PIDN:AAC46586.1; PID:g797299
  C:Genetics:
  A:Introns: 131/1, 222/2

alignment_scores:
  Quality: 182.50      Length: 650
  Ratio: 0.583         Gaps: 31
  Percent Similarity: 48.154  Percent Identity: 22.000

alignment_block:
  US-09-306-689-12 x A57036 ..
  Align seg 1/1 to: A57036 from: 1 to: 2491

412 TTRACAGATTGATGCAACGCGCCGAGAGTTGGGATTGAGTACAAAG 461
      |||||
1064 leuGlyAspLeuValIleSerAlaThrGlnSerAlaSerValaIaIa 1080
462 AGAAGAACCAATATATGCAACAGCTCAACCAAGCTTAGCGACGATTC 511
      |||||
1080 eSerAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 1095
512 AAACCGCTATT...GGCTTAAGTACAGCTGCGATTGTTATCCGCTCCA 558
      |||||
1095 InThrLysLeuMetGlyIleThrLysGlnLeuMetIle..... 1107
559 CAATATGATTAATGCTACAGAAACTAAAGCAGCCCAACGATTAGTTTC 608
      |||||
1108 AspMetGlnGlnLeuIleArg.....Al 1115
609 TGCGGAACCATTTGCAAAATGCAATTAAGCAACCAACGATTA..... 654
      |||||
1115 aSerArgSerValArgSerAsnProAsnAspArgArgSerGlnLeuLeu 1132
655 .....TCTGGCATTCATATTTTAGGCTCAGTATTTGGCTGGA 693
      |||||
1132 euAspArgArgSerAsnAspValSerIleSerThrAlaIaIaLeuValGly 1148
694 .....ATGCAATTTAGATGAGCGCTTACAGATTA 722
      |||||
1149 SerThrAlaAsnValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 1165
723 CAGCAACCAACATGCTCTGTAACGCTGAGCTGAGCTAACAAATTC. 771
      |||||
1165 eSerAsnLeuLeuSerLeuLysMetGlySerLeuGlnSerIleLeuSerG 1182
772 .....TTAATTAATAATTTCTAATTCAGTA 798
      |||||
1182 InProThrGlnGlnPheAlaPheTyrValGlnGlnIleAlaIaIaSerThr 1198

```

```

799 AAAACACTGACGAATTTGGTACCAAAATTAGCAATTTGGTCCAAACT 848
1199 LysAlaLeuSnaAlaAlaSerGlnGlnValValAlaMet...AlaArgAs 1214
849 ACAAAATATCAAAAGCTTAGGACTTTAGAGACAAACTCAAAATATCG 898
1214 nLysAsnLeuLysGlnLeuGlyAla..... 1222
899 GTGACACTGATAAGCTGGCCTGTTAGATGTTATCTCAGGCTATTA 948
1223 .....SerAlaLysIleThrAlaSerAlaLeuSerThrLeuVal 1235
949 TCGGGCCCAACAGCTGCATTTGACTT.....GCAGATAAAA 986
1236 SerAlaLysGlnAsnAlaIleValLeuThrGlnAsnGlnAlaThrLysAs 1252
987 TGCCTCAACAGCTAAAAAGTGGTGGGCTTTGAAATGGCAAAACCAAG 1036
1252 nAlaIleLeuAlaSerThrValAlaLeuGly.....GlyGlnI 1265
1037 TTGTTGGT..... 1044
1265 IleIleGlyLeuLeuAspPheSerLysAlaArgIleAlaAsnTyrLysAsp 1281
1045 .....AATATTACCAAAAGCCGTTCTCTTCTTA 1270
1282 ProIleTyrAspGlnAsnLeuIleAsnGlnAlaLysSerValGlnAspH 1298
1071 CATTTTACCCACAGCTGTGACAGCTTTATCTTCAACTGG..... 1113
1298 sLeuVal.....LysValGlyArgSerLeuGlyLysPylAsnAsn 1313
1114 .....CCTGGCTGCTTTAATGCTTCTACTGTTCTCTT 1149
1313 hrlIeCysAspGlnAlaValAlaAspArgIleIleGlnAlaThrArgSerLeu 1329
1150 .....GCAGATTAGCCCATTAAGATTGCGCGATTGCGCGATAATTAA 1193
1330 AspLysThrIleLeuPro...AspThrSerGlyLeuGlnThrAsnAlaH 1345
1194 T.....CATGCMAAAAGTTTAGAGAGTTATGCCGCAAGCTTTA 1231
1345 sLeuGlnMetLeuHisGlnGlnSerLeuLeuAlaIleThrGlnAlaSerL 1362
1232 AAAAATTAGCTATGACGAGATATTATATAGCAATATCAGCG... 1278
1362 yslLysLeuGlySerIleThrSerAsnLeuValAsnSerLysAsnAsnSer 1378
1279 .....GGACAGGAGACTATTGATCA..... 1299
1379 AspLeuValGlySerGlySerThrAspAlaGlnAlaArgIleIleGlnMetI 1395
1300 .....TCGCTTACTGCAATTAATACCG 1321
1395 eGlnAlaAlaLysHisValValAlaHisCysSerIleSerThrTyrAsnProA 1412
1322 CATGGCCGCTATGCTGTGCTGTGCTGTGCTGCAGCGATTAACA 1371
1412 spIleLeuLeuProAlaLysSerIleLeuAspAlaSerGlnMetLeuThr 1428
1372 TTTGAAAAAGTTAAACATATCTTGTATCATCAGAAATAGCAAAAAAGAA 1421
1429 AlaAsnGlnAlaAspValAlaAsnHisValLeuSerHisAla..... 1441
1422 AGTACCATTCMAAACTGGTCCAGAGAGCTGATTTGCTAAAGAGTGC 1441
1441 ..... 1441
1472 CTAATTTAAAGCAACTAAAGATGAGAAATCGAAGAAATCATCGGTCAA 1521
1443 .....AlaThrIleAlaAlaLacysThrGlnGlnLeuLeuGlyIle 1454
1522 AATGGCAGGCGGATCACC.....TCAAAAGCAAGTGTATGATCTTATCGC 1565

```

```

1455 ThrArgGlnArgAlaSerGlnPheAsnGlnGlnAspGlnGlnGlnValG 1471
1566 AAAAGTAAACCGCAAAATTAACCCAGATGAGCTATCAAAAGTTGTGATA 1615
1471 nValArgAspGly.....IleValLys 1479
1616 ACTATGAAATTCCTCAACATATGCAAAATATGCAAAACAGCTTGATTAAG 1665
1479 erThrGlnGlnLeuAlaHisAla.....ThrSerSerLeuAlaArg 1492
1666 TTAATCTCATCTGTAAGTGCATTTACCTGCTTAATGATTCGAGAAATGT 1715
1493 AlavalysSerVal.....ThrSerLysGlnProGlyAlaLysAl 1506
1716 ATTAGTGCTCCCAACTTCATGTTGATGCAAAAGTTTATCTTCTTCAT 1765
1506 aMetIleSer...GlnSerLeuLysAspLeuGlnSerAla...IleAsnAsn 1521
1766 TTGCTAGGAGATCTCAGCATGAGCTAGC.....GCCTGCCCTGAGC 1809
1522 LeuLeuIleThrSerSerValProAlaSerGlnArgGlyIleGlyIleAl 1538
1810 AGCGGTTCTCAAGATTGAGCTAGCGCTGCGTCCGGTGGCTCTAGCCA 1859
1538 aAspPheAsnLysLeuMetSerThr...CysArgSerValSerThrAlas 1554
1860 GC.....ATTGAGCTACGCGCTGC..... 1879
1554 erSerGlnLeuIleIleSerAlaSerLysSerGlnLysProLysAsp 1570
1880 .....GC 1881
1571 IleGlnLeuSerSerIleLeuSerGlnAsnAlaValLeuMetThrAsnSe 1587
1882 CTGCGACAGCGGTGACCAAGATTGAGCTAGCGCTGCGTCCGGTGCATC 1931
1587 rLeuLysAspIleIleLysValThrSerSerMetMetProGlyValAsn 1604
1932 TCACATTTGAGCTAGCGGCTGCGCTGCGAGCGGCTTCCTCAAGATGGA 1981
1604 heCysGlnGlnAlaIleGlnIleAlaGlnArgAlaIleSerAspLeuSer 1620
1982 GCTACGGCTGCTGCTCCGCGTGTCTTAGCCAGCATTTGAGCTAGCGCC 2029
1621 SerValAla...LeuSerValAlaValGlySerPheAspSerSerAla 1635

seq_name: p1r1:KRH07

seq_documentation_block:
  keratin I, type II, cytoskeletal - human
  N:Alternate names: 67k type II epidermal keratin; cytokeratin 1
  C:Species: Homo sapiens (man)
  C>Date: 04-Dec-1986 #sequence-revision 22-Oct-1999 #extLchange 10-Dec-1999
  C:Accession: A22940; A02950; A43342
  R:Johnson, L.D.; Idler, W.W.; Zhou, X.M.; Roop, D.R.; Steinert, P.M.
  Proc. Natl. Acad. Sci. U.S.A. 82, 1896-1900, 1985
  A:Reference number: A22940; MUID:85166239
  A:Accession: A22940
  A:Molecule type: DNA
  A:Residues: 1-643 <OH>
  A:Cross-references: GB:M98776; GB:M11215; GB:M11845; GB:M11846; NID:g1843461
  A:Note: translation of initiator Met is not shown
  R:Steinert, P.M.; Parry, D.A.D.; Idler, W.W.; Johnson, L.D.; Steven, A.C.; Roop, D.R.
  J. Biol. Chem. 260, 7142-7149, 1985
  A:Title: Amino acid sequences of mouse and human epidermal type II keratins of M-r 67
  late filament subunits.
  A:Reference number: A92535; MUID:85207740
  A:Accession: A02950
  A:Molecule type: mRNA
  A:Residues: 151-183,'K',185-199,'M',201-204,'K',206-236,'S',238-239,'R',241-356,'Y',3
  'S',638-643 <STE>
  A:Cross-references: GB:M10938; NID:g186787; PIDN:AAA36153.1; PID:g386854

```



```

1177 ATGGCCGATAAATTAAATCATGCAAAAAGTTTAGAGAGTTATGCCGACG 1226
      :::::|||||
367 .....GluAlaGluSerLeu.....TyrGlnSerIly 375
      :::::|||||
1227 CTTTAAAAAATTA.....GGCTATGACGAGATTAATTATTATAG 1264
      :::::|||||
375 sTYrGlnGlnLeuGlnIleThrAlaGlnArgHisGlyAspSerVal.... 390
      :::::|||||
1265 CAGAAATATCAGCGGGGCAACAGGAGCTATTGATGCATCGGTTACTGCAATT 1314
      :::::|||||
391 .....ArgHisSerIlyIleGlnIleSerGlnLeu 400
      :::::|||||
1315 AATACCCGATTGGCGGCTATTGCTGTGGTGTGTCTGCTGTCGACGCCGA 1364
      :::::|||||
401 AsnArgValIleGlnArgLeuArgSer..... 409
      :::::|||||
1365 TTTTACATTGAAAAAGTTAAACATTAATCTTGTCATCAGCAATAGCAAAA 1414
      :::::|||||
410 .....GluIleAspHisValIlySL 416
      :::::|||||
1415 AAGAGAAAGTACCATTCAAAACTGTTCCGAGAGGCTGATTTTGTAAA 1464
      :::::|||||
416 yscHnIleSerAsnLeuGlnGlnSerIleSerAspAlaGlnArgGly 432
      :::::|||||
1465 GAAGTGCCTAATTATAAGCAACTAAAGATGAGAAA..... 1500
      :::::|||||
433 Gln.....AsnAlaLeuLysAspAlaLysAsnLysLeuAsnAs 445
      :::::|||||
1501 .ATGGAAGAAATCATCGCTCAAAATGGCGAGCGGATCACCTCAAGACAG 1549
      :::::|||||
445 pLeuGlnAspAlaLeuGlnGln..... 452
      :::::|||||
1550 TTGATGATCTTATCGCAAAAGGTAAACGCAAAATTACCAAGATGAGCTA 1599
      :::::|||||
453 .....AlaIlySGlnAspLeu 457
      :::::|||||
1600 TCAAAAGTTGTGATTAAGTATGAAATGCTCAAAACATAGCAAAAATGTGAC 1649
      :::::|||||
458 ThrArgLeuLeuArgAspTyrGlnGlnIleuMetAsnThrIys..... 471
      :::::|||||
1650 AAACAGCTTGAATGAATTCATCTGTAAGTCGATTAACCTGCTGCTA 1699
      :::::|||||
472 .LeuAlaLeuAspLeuGlnIleAlaIleThrTyrArgThrLeuLeuGlnIlyG 488
      :::::|||||
1700 ATGATTTAGAGA...AATGTATTAGTGCCTCAACTCAATGCTGATCAAA 1746
      :::::|||||
488 LngIuSerAlaGlnMetSerGlyGlnLysAlaProAsnValSerValThrVal 504
      :::::|||||
1747 AGTTTATCTCTCTTCAATTGTCTAGGGGATCTGACATTTGAGAGCTACGG 1796
      :::::|||||
505 SerThrSerHisThrSerIleSerGlyGlyIlySerArgGlyGly 521
      :::::|||||
1797 CTTGCGCCCTGGCGGCTGCTCTCAAGATGG.....AGCT 1831
      :::::|||||
521 yGlyGlyTyrIlySerIlySerIlySerTyrGlySerGlyIlySerT 538
      :::::|||||
1832 ACCGCTCGGCTCGGCTGCTAGCCAGCATTTGAGCTAGCGGCTCGCGC 1881
      :::::|||||
538 yrgIlySerGlyGlyIlyGlyIlyGlyIlyArgGlySerTyrGlySerGly 554
      :::::|||||
1882 CTTGCGCAGC...GGTACCAAGATTTGAGCTAGCGGCTGCGCTCGGGTGG 1928
      :::::|||||
555 GLyIlySerTyrGlySerGlyIlySerSerTyrIlySerGlyIlyGly 571
      :::::|||||
1929 ATCTCAGCATTTGAGCTAGCGGCTCGGCTCGGAGGCTTCTCAAGATT 1978
      :::::|||||
571 yGlyGlyHisIlySerTyr.....GlySerGlySerSerSerG 584
      :::::|||||
1979 GAGAGCTAGCGCTCGGCTCGGAGTCTTAGCCAGCATTTGAGAGCTAGCGC 2028
      :::::|||||
584 IyGlyTyrAlaGlyGlyIlySerGlyIlyGlyIlyGlyIlySerSerGly 600
      :::::|||||
2029 CTGCGCCCTGGCGAGCGGTAGCCAAAGATTGAGCTAGCGGCTGCTCGGG 2078

```

```

601 ArgGlySerGlyGlyIlySerSerGlyIlySerSerGlyIlyArgIlyse 617
2079 TGGATTC 2085
617 rSerSer 619
seq_name: p1r2:C83339
seq_documentation block:
hypothetical protein PA2462 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000
C:Accession: C83339
R:Stover, C. K.; Pham, X. O.; Erwin, A. L.; Mizoguchi, S. D.; Warrenner, P.; Hickey, M. J.;
adman, S.; Yuan, Y.; Brody, L. L.; Coulter, S. N.; Folger, K. R.; Kas, A.; Lardig, K.; L.
.; Lory, S.; Olson, M. V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950
A:Accession: C83339
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5627 <STC>
A:Cross-references: GB:AE004673; GB:AE004091; NID:g9948501; PIDN:AA05850.1, GSPDB:GN
A:Experimental source: strain PA01
A:Genetics:
A:Gene: PA2462
alignment_scores:
Quality: 171.50 Length: 725
Ratio: 0.482 Gaps: 42
Percent Similarity: 49.103 Percent Identity: 22.483
alignment_block:
US-09-306-689-12 x C83339 ..
Align seg 1/1 to: C83339 from: 1 to: 5627
31 TGGAGCTAGCGGCTCGGCTCGGAGCGGTTCTCAAAATTTGAGACTACGG 80
|||||:::|||||:::|||||:::|||||
5018 TrpIlySerIser.....AsnSerGly.LeuGlyValGlyAlaIlnL 5031
81 CTTGCGCTCGGCGGCTGCTAGCC.....AGCATTTGAGCTAGCGGCTGC 124
|||||:::|||||:::|||||:::|||||
5031 euAlaIleAlaIleValAlaIleTyrPheThrAlaGlyAlaIleAlaSerAla 5047
125 GCCCTGGCAGCGGTAGCCAAAGATTGAGCTAGCGGCTGCGGTGGA 174
|||||:::|||||:::|||||
5048 AlaLeuGlySerMetAlaGlyValGlyAla..... 5057
175 TCTAGCATTTGAGACTAGCGGCTCGGCTCGGCGGCTGCTCAAGATTG 224
|||||:::|||||:::|||||
5058 .....GlySerGlySerMetMetAlaAlaAla.....G 5067
225 GAGCTAGCGGCTCGGCTCGGCTGCTAGCCAGCATTTGAGACTAGCGGC 274
|||||:::|||||:::|||||:::|||||
5067 IySerThrAlaMetValGlnAlaGlyThrAla...ValGlyThrAlaAla 5082
275 TGGCGCTCGGCGAGCGGTAGCCAAAGATTGAGCTAGCGGCTGCGGGT 324
|||||:::|||||:::|||||:::|||||
5083 .....AlaGlyTyrPalaAsnAlaIleAlaGlyThrAla...ValAlaMe 5095
325 GGAATCTAGCTTCCCAAAAAGTGGGGCAAAAAAATTATCTCTATATTC 374
|||||:::|||||:::|||||:::|||||
5095 tGlyMetAlaSer.....AsnGlyAlaIleAs 5104
375 CCAAAATTACCAATAT.....GATACGAA 399
|||||:::|||||:::|||||
5104 ertThrIleAsnAsnArgGlyAsnLeuGlyAspValIlyAspAspValThr 5120
400 CAAAGTAATGTTTACAGATTAGTCAAAAGCGGCGAAGAGTTGGGAT 449

```

```
5121 SerSerAspAlaLeuArgLysGlyTyrValValAlaGlyThrThrAlaGlyLeu 5137
      :::::::::::::::::::: ||| :::::::::::::: ::::
450 T.....GAGGTCAAAAGAGAGACAGCAATTAATATTGCA 484
      :::::::::::::::::::: :::: :::::::::::::: ::::
5137 ThrThrAlaGlyValTyrAspLysTrpThrSerThrGlnThrGlyThrSerT 5154
      :::::::::::::::::::: :::: :::::::::::::: ::::
485 CAGGTCCAACACGTTTAGCGACGATTCACACCGCATATGGCTTA...ACT 531
      ||||| :::: :::::::::::::: :::: ||||| |||||
5134 ThrAlaLeuProfsnThrGlyAlaValAlaProAlaAlaGlyLeuGlyThr 5170
      :::::::::::::::::::: :::: :::::::::::::: ::::
532 GAGCGTCGATGTGTATCCGCTCCACAAATTTGATTAATGCTACAGAA 581
      :::::::::::::::::::: :::: :::::::::::::: ::::
5171 TrpGlnGlyVal..... 5174
      :::::::::::::::::::: :::: :::::::::::::: ::::
582 AACTTAAGACGCCCAAGCATTAAGTTCTGCCCAAGACATTTACAAATG 631
      ||||| :::: :::::::::::::: :::: :::::::::::::: ::::
5175 .....GlyGlnPhe.....ThrSerAsnGlnLeuLeuGlnAsn. 5185
      :::::::::::::::::::: :::: :::::::::::::: ::::
632 CAATTAAGCCAAACCTGATTAATCTGCATCATCTATTAGGCTCA 681
      ||| :::: :::::::::::::: ::::
5186 .....GlyThrSerValLeuLeuAspArg 5193
      :::::::::::::::::::: :::: :::::::::::::: ::::
682 GTATTGGCTGGAATG...GATTAGATGAGCGCTTACAGATAACAGCAA 728
      :::::::::::::::::::: :::: ||||| :::::::::::::: ::::
5194 AlaLeuGlyGlyLysGlySerLeuGlyAspAlaLeuGlnAsn..... 5207
      :::::::::::::::::::: :::: :::::::::::::: ::::
729 CCACATGCTCTTGTCTAAAGCTTGCTGAGCTTAACAAATTCATTAAATG 778
      :::::::::::::::::::: :::: :::::::::::::: ::::
5207 ..... 5207
      :::::::::::::::::::: :::: :::::::::::::: ::::
779 AAAATATGCTTAATCAGTAAAAACACTGACGAATTTGGTAGCAAAAT 828
      :::::::::::::::::::: :::: :::::::::::::: ::::
5208 ..SerLeuAlaAsnAla.....Phe 5213
      :::::::::::::::::::: :::: :::::::::::::: ::::
829 AGTCAATTTGGTTCAAACTCAAAATATCAAAAGCTTAGGACCTTAG 878
      :::::::::::::::::::: :::: ||||| :::::::::::::: ::::
5214 AlaAlaTyrGlyPheLysLeu.....IleGly 5222
      :::::::::::::::::::: :::: :::::::::::::: ::::
879 AGACAACCTCAAAATATTCGCTGAGCTTGAATGAACCTGCGCTTGGTTAG 928
      |||| :::: :::::::::::::: :::: ||||| :::::::::::::: ::::
5222 TyrAspThrTrpHis.....GlyValLeuAspAspLysSerLeuGlyLysI 5237
      :::::::::::::::::::: :::: :::::::::::::: ::::
929 ATGTTATCTCAGGCGCTATTATCGGCGCAACAGCTGCACTGTACTGCA 978
      :::::::::::::::::::: :::: :::::::::::::: ::::
5237 LeGlyLeuHisAlaLeuMetGlyGlyLeuAlaAlaGlyValGlyGly 5253
      :::::::::::::::::::: :::: :::::::::::::: ::::
979 GATTAATAATGCTTCAACAGCTAAAAAAGGGTGGGCTTTGGAATTTGCG 1028
      ||| :::::::::::::::::::: :::: :::::::::::::: ::::
5254 AspPheAlaGlyThrGlyAla.....LeuAlaAlaGlyValAlaAsnGlyAl 5267
      :::::::::::::::::::: :::: :::::::::::::: ::::
1029 AAACCAAGTTGTTGGTAATTAATTAACAAACCGTTCTTCTTACATTTTAG 1078
      | :::::::::::::::::::: :::: :::::::::::::: ::::
5267 a.....LeuValAspSerLeuAlaGlyGlnThrAlaSerLeuProIle 5282
      :::::::::::::::::::: :::: :::::::::::::: ::::
1079 CCCAAGCGTTGTCACAGCTTATCTCACTGCGCGCTGCGCTGCTTGA 1128
      :::::::::::::::::::: :::: :::::::::::::: ::::
5282 spAspLysLysGlyLeuLeuIleMetSerSerGlnLeuIleGlyValLeu 5298
      :::::::::::::::::::: :::: :::::::::::::: ::::
1129 ATTGCTTACGTTCTTCTCTGCGATTAAGCCATTAGCA..... 1167
      ||||| :::: :::::::::::::: ::::
5299 AlaAlaSerThrGlnGlyAspAlaAspAlaLysSerLeuGlnThrGlyAl 5315
      :::::::::::::::::::: :::: :::::::::::::: ::::
1168 ...TTTGCCGCTATGCGCGATTAATTAATCAATGCAAAAGTTTAGACA 1213
      ||||| |||| :::: :::::::::::::: ::::
5315 aspValAlaGlyAsnAlaThrGlnHisAsnTyr.....LeuSerH 5329
      :::::::::::::::::::: :::: :::::::::::::: ::::
1214 GTTAATGCCGACGCTTAAAAAATTAAGCTATGAGGGA.....CATAT 1257
      :::: ||||| :::: |||||
5329 IstTrpGlnGlnGlyLysLysArgGlnGlyValAspLysLysAspLys 5345
      :::: ||||| :::: |||||
1258 TTATTAGCAGATTAATCAGCGGGAACAGGCACTATTGATGACATCGCTTAC 1307
      ||| :::: :::: ||||| ||||| :::: ::::
```

```
5346 GlnLeuCysLys.....ThrGly...IleGlyAlaLysTrpAl 5357
      :::::::::::::::::::: :::: :::::::::::::: ::::
1308 TGCATTAATTACC.....GCATTGGCCGCTATGCTGGTG 1342
      :::::::::::::::::::: :::: :::::::::::::: ::::
5357 AlIleIleSerAlaGlnGlnAspValGlyTyrLeuValGlyValGlyGlyG 5374
      :::::::::::::::::::: :::: :::::::::::::: ::::
1343 GGTGTCTGTGCTGCGACCGCATTTA.....ACATTGAAAAAGTT 1383
      ||| :::::::::::::::::::: :::: :::::::::::::: ::::
5374 TyrIleGlyLeuSerThrAlaGlnThrAlaValGlyValTyrGlnLeuVal 5390
      :::::::::::::::::::: :::: :::::::::::::: ::::
1384 AAACATTAATCTTGTATCAGCAATAGCAAAAGAGAAAGTACCATTTCA 1433
      ||| ||| :::: :::::::::::::: ::::
5391 Lys...AsnTrpArgGlnThrTyrAlaAlaLeuGlnGlnLeuAlaThrSe 5406
      :::::::::::::::::::: :::: :::::::::::::: ::::
1434 AACCTGTTCCGAGAGCT.....GATTGCTTAAGACGTGCTTA 1474
      :::::::::::::::::::: :::: :::::::::::::: ::::
5406 rProGlnPheArgGlnGlnPheGlyAspAsnTyrLeuLysGlyLeuGlnG 5423
      :::::::::::::::::::: :::: :::::::::::::: ::::
1475 ATTATTAAGCACTAAAGATGAGAAAAATGAGAA..... 1509
      :::::::::::::::::::: :::: :::::::::::::: ::::
5423 LuArgAlaAlaPheLeuThrGlnAlaTyrGlnAspAlaGlyTyrGlnGly 5439
      :::: ||| :::: :::::::::::::: ::::
1510 .....ATCATCGTCAAAATGCGCGAGCGATCACCTCAACCAAGTTGA 1553
      ||| :::: ||| :::: ||||| :::::::::::::: ::::
5440 SerValThrAlaGlyValGlnGlyGlyArgPheAlaAlaGlyLeuValGly 5456
      :::::::::::::::::::: :::: :::::::::::::: ::::
1554 TGAATCTTATGCA...AAAGTAAACGGCAAAATTAACCAAGATGAGCTAT 1600
      ||| ||| ||||| :::::::::::::: ::::
5456 yValLeuThrAlaValLysGlyGlyAlaGlnIleThrAla..... 5469
      :::::::::::::::::::: :::: :::::::::::::: ::::
1601 CAAAAGTCTTGAATACATGATGAAATGCTCAACATAGCAAAATGAGCA 1650
      :::: ||| :::: ||||| :::::::::::::: ::::
5470 .....LysLeuProThrAlaAlaLysAsnLeuVal 5479
      :::::::::::::::::::: :::: :::::::::::::: ::::
1651 AACAGCTTAGATAAGTAAATCATCTGATCTGATGCAATTTACCTGCTTAA 1700
      ||||| :::: :::: ||| :::: ||||| :::::::::::::: ::::
5480 AsnAlaIle.....AlaGlnSerProValSerGlySerIle 5491
      :::::::::::::::::::: :::: :::::::::::::: ::::
1701 TGAATTCAGCAAAATGTAATAGTGGCTCCAACTCAATGTTG.....G 1741
      :::: ||||| :::: ||||| :::: |||
5491 tSerSerGlnLeuGlyAlaValGlyAspLeuGlyArgLeuGlyGlyG 5508
      :::::::::::::::::::: :::: :::::::::::::: ::::
1742 ATCAAAAGTTATCTCTCTTCAATTTGCTAGGGGATTCAGCATTTGAGAC 1791
      :::::::::::::::::::: :::: :::::::::::::: ::::
5508 TyrLysGlyTyrValAspIleLeuSerHisGlnAlaLysGlnHisIleLeu 5524
      :::::::::::::::::::: :::: :::::::::::::: ::::
1792 TAGGCGCTCGCCCTGCGAGCGGTTCTCAAGATTGGAGCTAGCGCTGCG 1841
      ||||| :::: ||||| :::: ||||| :::::::::::::: ::::
5525 TyrGlyAspLysProGlySerGlyGlnHisLeuTrp..... 5536
      :::::::::::::::::::: :::: :::::::::::::: ::::
1842 TCCGGGTGCTCTAGC.....CAGCATTTGAGCTACGGCC 1876
      ||||| :::: :::: ||||| :::: |||||
5537 .ProGlyGlnAlaGlyLysThrValPheProGlnAsnTrpSer .AlaAsp 5552
      :::::::::::::::::::: :::: :::::::::::::: ::::
1877 TGCGCCCTGCGACCGGTACCCAAAGTTGAGCTAGCGCTGCGTCCGGGT 1926
      :::: ||||| :::: ||||| :::: |||||
5553 LysIleValHisGlnValGlyAspIleAlaThrSer .ProSerThrLysT 5569
      :::: ||||| :::: ||||| :::: |||||
1927 GG.....ATCTCAGCATTTGAGCTAGCGGCT 1952
      :::: :::: :::: |||
5569 rPtyAlaGlnThrGlyThrGlyGlyValTyrThrSerLysGlyAspPro 5585
      :::: ||||| :::: ||||| :::: |||||
1953 GCGCCCTGCGACGGTTCGA.....AGATTGGA 1981
      ||| ||| :::: :::: |||||
5586 AlaLysTrpValAlaTyrGlnValAlaArgAspGlyValArgMetArgVal 5602
      :::: ||||| :::: |||||
1982 GCTACGCGCTGCGTCCGG 2000
      :::: ||||| :::: |||||
5602 TyrGlnProAlaThrGly 5608
      :::: ||||| :::: |||||
```

seq_name: p1r2:A75564

seq_documentation_block:

conserved hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #ext_change 17-Mar-2000
 C:Accession: A75564
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M., Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioreistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; M01D:20036896
 A:Accession: A75564
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1467 <MB1>
 A:Cross-references: GB:AE001870; GB:AE000513; NID:g6457732; PIDN:AAF09665.1; PID:g645773
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR0075
 A:Map position: 1

alignment_scores:
 Quality: 169.00 Length: 838
 Ratio: 0.426 Gaps: 40
 Percent Similarity: 47.375 Percent Identity: 22.196

alignment_block:
 US-09-306-689-12 x A75564 ..

Align seg 1/1 to: A75564 from: 1 to: 1467

```

43 CTGCGCCCTGCGACGGCTTCT...CAAGATTGAGCTACGCGCTGCTCC 89
||||| ||||| ::| ||||| ::| |||||
194 LeuArgspgLyAlaThrAspLeuGlnGlnGlyThrValLysLeuArgAs 210
90 GCGTGGCTCT.....AGCCAGCATTTGAGCT 115
||||| ||||| ::| |||||
210 pGlyAlaSerAspLeuAlaAspGlyAlaAsnArgAlaAlaAspLysSerA 227
116 ACGGCTGCGCCCTGCGACGGCTACGCAAGATTGAGCTACGCGCTGCT 165
||| ||||| ||||| ||||| |||||
227 rGlySerLeuAlaAspGlyAlaGlyThrLeuSerSerGlyValGlyArgLeu 243
166 CCGGGTGATCTCAGATTGAGCTACGCGCTGCGCCCTGCGACGGCTTC 215
||||| ||||| ||||| ||||| |||||
244 ThrAspGlyThrValLysLeuSerAspGlyLeuArgGlnLeuGluAlaAl 260
216 TCAAGATTGAGCTACGCGCTGCTCCG.....GGTGGCTCTAGCCAGC 259
::| ||| ||| ||| |||
260 AlaProGlyGlnThrGlnLeuAlaProLeuArgSerGlyAlaGlyGln. 276
260 ATTGAGCTACGCGCTGCGCCCTGCGACG..... 288
||||| ||||| |||||
277 .....LeuArgThrGlySerGlnSerLeuAlaSerGlyLeu 288
289 .....GGTAGCCAAGATTGAGCTACGCGCTGCGCTCCG. 321
||||| ||||| ||||| ||||| |||||
289 GlyGlnLeuAlaAspGlyThrAspGlnLeuAlaGlnGlyAlaLysLysAl 305
322 ....GGTGGATCTAGCTTCCCAAAACTGGGGCAAAAAAATTTCCTCT 367
||||| ||||| ||||| ||||| |||||
305 AsnGlyGlyAlaGlnGlnValAlaAlaGlnGlyAsnLysGlnLeuAlaAlaG 322
368 ATATTCCCAAAATTACCAATATGATACGAAACAAGTAATGTTTACAG 417
::| ||||| ||||| ||||| |||||
322 LnuLeuProGln.....LeuGlnSerGlyLeuThr 331
418 GATTTACTCAAGCGCGCAAGAGCTGGGATTTAGGTACAAAGAGAAGA 467
::| ||||| ||||| ||||| |||||
332 GlnLeuSerSerGlyAlaAspLysLeuAla..... 341
468 ACGCAATATATTGCAACAGCTCAACACAGTTTGGCAGCATTTCAACCG 517
||| ||||| ||||| |||||

```

```

342 .....AlaGlyAlaLysSerAlaAsnGlyAlaGlnGlnInv 354
518 CTATTGGC.....TTACGTAGCGT..... 537
::| ||| ||||| |||||
354 AlaAlaGlyGlyAsnLysGlnLeuAlaGlnGlnLeuProLysLeuGlnGly 370
538 GGCATTGTGTTATCCGCTCCCAAAATTGATAAATGCTACAGAAACTAA 587
||| ||||| ||||| ||||| |||||
371 GlyLeuThrGlnLeuSerSerGlnGlyAlaAspLysLeuAlaGlnGlyAla 387
588 AGCAGGCCAAGCATTTAGTTCTGCCGAAGCATTTGTCAAAATGCAATA 637
||| ||| ||||| ||||| ||||| |||||
387 sSer.....AlaAsnGlyGlyAlaGlnGlnValAlaGlyGlyAsnLysG 402
638 AAGCAAAACTGTATTATTCGCAATTCACCTATTTAGCGCTGATTTG 687
::| ||| ||||| ||||| |||||
402 LnuLeuAlaAspGlnLeuProGlnLeuGlnSerGlyLeuThrGlnLeuSer 418
688 GCTGAATGATTTAGATGAGCCCTTACAGATTAACAGCAACCAACATTC 737
||| ||| ||||| ||||| |||||
419 GlnGlyAlaAspLysLeuAlaAlaGlyAlaLysSerAlaAsnGlyGlyAl 435
738 T.....CTTGCTAAAGCTGGCTTGAGCTTACAAATTCATTA..... 774
||| ||||| ||||| ||||| ||||| |||||
435 aGlnGlnValAlaAlaGlnGlyAsnLysGlnLeuAlaAspGlnLeuProGln 452
775 .....ATTGAAATATTGCTAAT 792
452 eugInSerGlyLeuGlyGlnLeuSerGlnGlyAlaAspArgLeuAlaSer 468
793 TCAGTAAAAACACTTGACGAATTTGTGAGCAAAATTAAT 831
||| ||||| ||||| ||||| |||||
469 GlyAlaLysAspAlaAsnAlaGlyAlaGlnGlnValAlaAlaGlyAsnLy 485
832 CAATTGTGTTCAAACTACAAATATATCAAA...GGCTTAGGCACTTAG 877
||| ||||| ||||| ||||| ||||| |||||
485 sGlnLeuAlaAlaGlnLeuProGlnLeuGlnSerGlyLeuGlyGlnLeu 502
878 GAGCAAACTCAAAATATATC...GGTGACTTGATAAGCTGGCTGCTGT 924
||| ||||| ||||| ||||| |||||
502 eugInGlyAlaAspArgLeuAlaSerGlyAlaLysGlyAlaAsnAlaGly 518
925 TTAGATTTATCTCAGG..... 942
519 AlaGlnGlnValAlaAlaGlyAsnLysGlnLeuAlaThrGlnLeuProGln 535
943 CTATTATCGGGCCGCAACAGCTGCATTTGCTGAGATAAAATCTT 991
||| ||||| ||||| ||||| |||||
535 nLeuGlnSerGlyLeuGlnLeuSerGlnGlyAlaAspArgLeuAlaAs 552
992 CAACAGCTAAAAAAGTGCGGCGGTTT..... 1020
||| ||||| ||||| |||||
552 eugAlaLysAspAlaAsnAlaGlyAlaGlnGlnValAlaAlaGlyAsn 568
1021 ...GAATTGGCAACCAAGTT.....GTTGGTAATAT 1049
||| ||||| ||||| ||||| |||||
569 LysGlnLeuAlaThrGlnLeuProGlnLeuGlnSerGlyLeuGlyGlnLe 585
1050 TACCAAAAGCCGTTCTTCTTACATTTAGCCCAAGCTGTGCAGCAGTT 1099
||| ||||| ||||| ||||| |||||
585 uSerGlnGly.....AlaAspArgLeuAlaSerGlyA 596
1100 TATCTTCACTGAGCGCTGTGCTGCTTATTTAGTTCTTACTGTTTCTCT 1149
||| ||||| ||||| ||||| |||||
596 LAsnAspAlaAsnSerGlyAlaAlaLysGlnLeuAla..... 607
1150 GCGATTAGCCATTAGCATTTGCCGCT.....ATTGCCGATAAAT 1190
||| ||||| ||||| ||||| |||||
608 .....AlaGlyAsnLysGlnLeuAlaAspGlnLe 617
1191 TAATCATGCAAAAAGT.....TTAGAGAGTTATGCCGAAGCT 1228
||| ||||| ||||| ||||| |||||
617 uProGlnLeuGlnSerGlyLeuGlnGlnLeuSerGlnGlyAlaAspArgL 634

```

```

1229 TTTAAATAATTAGCTATGAC..... 1248
      :: :: :: |||
634 euAlaSerGlyAlaAsnAspAlaAsnAlaGlyAlaAlaGlnIleAlaAla 650
1249 GGAGATAATTATTATAGCA.....GAAATACAGCGGGGAACAGG 1286
      ||| ::| |||
651 GlyAsnArgGlnLeuAlaThrGlnLeuProGlnLeuGlnSerGlyLeuG 667
1287 GACTATTGATCGATCGGTTACTGCAATTATACCGGATTCGCCGCTATTC 1336
      ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
667 nGlnAlaSerAlaGlyAlaAspArgLeuAsnThrGlyAlaAlaGlnLeu 684
1337 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1386
      || ||| ::| ::| ::| ::| ::| ::| ::| ::| ::|
684 IaThrGlyThrLeuGlnLeuAsn.....AspAlaValLys 695
1387 CATTAATCTTGTCATGCAAGATAGCAAAAAAGAAAGTACCATTCATAA 1436
      ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
696 AsnGlnAlaLeuLeuProGlyGlnLeuLysGlnGlyAlaAlaValLys 712
1437 CTGGTTCGAGAGCGCTGATTTGCTAAAGAACTGCCATTTATTAAGCAA 1486
      | ::| ||| ||| ::| ::|
712 n.....SerGlyAlaMetGlyValHisSerGlyThrAsnT 724
1487 CTAAAGATGAGAAATCGAAGAAATCATCGTCATAAATGCGCGAGCAGTC 1536
      || ::| ::| ::| ::| ::| ::| ::| ::| ::|
724 hLeuSerAsnAlaLeuProGlnLeuGlnAlaGlyAlaGlyValAlaVal 740
1537 ACCTCA.....AAGCAAGTTGATGATCTTATCGCAAAAGGTAA 1574
      ||| ::| ::| ::| ::| ::| ::| ::| ::| ::|
741 ThrGlyAlaSerGlnLeuAlaGlnGlyAlaAspLysLeuAlaAlaGly 757
1575 CGGCAAAATTTACCCCAAGATGACCTATCAAAAGTTGTTGTAACCTGAAT 1624
      ::| ::| ::| ::| ::| ::| ::| ::| ::|
757 rAlaGlnLeu..... 760
1625 TCCTCAAAACATAGCAAAATGTGACAACACGCTTAGATAAGTTA..... 1668
      ||| ::| ::| ::| ::| ::| ::| ::| ::|
761 .....SerGlnGlyAlaThrThrLeuSerSerLysLeuGlu 773
1669 .....ATCATCTGTGATGATGATGATGATGATGATGATGATGATGAT 1703
      ::| ::| ::| ::| ::| ::| ::| ::| ::|
774 AlaGlyGlnGlySerThrAlaAlaValGlnGlyValGlnGlnLeuSer 790
1704 TTCGGAATGTAATTAGTGGCTCCAACTCAATGTTGATCA..... 1746
      ::| ::| ::| ::| ::| ::| ::| ::| ::|
790 uGlyAlaThrArgLeuAlaAlaGlyThrAlaGlnLeuSerGlnGlyAlaVal 807
1747 ...AGTTTACTCT..... 1758
      ::| ::| ::| ::| ::| ::| ::| ::| ::|
807 aLThrLeuSerSerLysLeuLysGlnAlaGlnGlnGlyThrThrAlaAla 823
1759 CTTCAATTTGTAGGAGATCTCAGCATTTGAGCTACGCGCTCGCC..... 1804
      ::| ::| ::| ::| ::| ::| ::| ::| ::|
824 ValGlnGlyAlaGlnGlnLeuGlyGlnGlyAlaAspLysLeuAlaAlaG 840
1805 .....CTGGCAGCGGTTCTCAAG 1822
      ||| ::| ::| ::| ::| ::| ::| ::| ::|
840 lYThrAlaGlnLeuSerGlnGlyAlaAsnThrLeuSerSerLysLeuLys 856
1823 .....ATTGAGCTACGCGCGCGCGCGCGCGCTGTCAGCCAGCAT 1863
      ||| ::| ::| ::| ::| ::| ::| ::| ::|
857 GluAlaGlnGlnGlyThrThrAlaAlaValGlnGlyAla.....GlnGln 872
1864 TCGAGCTACGCGCTCGCGCTCGAGCG..... 1891
      ||| ::| ::| ::| ::| ::| ::| ::| ::|
872 uGlyGlnGlyAlaAspLysLeuAlaAlaGlyThrAlaGlnLeuSerGlnG 889
1892 .....GTAGCCAGATTGGAGCTACG 1912
      ||| ::| ::| ::| ::| ::| ::| ::| ::|
889 lYAlaThrThrLeuSerSerLysLeuLysGlnAlaGlnGlnGlyThrThr 905

```

```

1913 GCCTGCGTCCGGGTGATCTCAGCATTTGGAGCTACGCGCTCGCCCTGCGC 1962
      ||| ::| ::| ::| ::| ::| ::| ::| ::|
906 AlaAlaValAlaGlnGlyAlaGlnGlnLeuGlnGlnGlyGlnAlaAspLysLeuAla 922
1963 A.....GCGGTTCTCAAGATTGGAGCTACGCGCTCGCTCGCGGTGG 2003
      | ::| ||| ::| ::| ::| ::| ::| ::| ::| ::|
922 aAlaGlyThrAlaGlnLeuSerGlnGlyAlaThrThr...LeuSerThrL 938
2004 CTTCTAGCCAGCATTTGA...GCTACGCGCTCGCGCTCGACAGCGTACGC 2050
      ||| ::| ::| ::| ::| ::| ::| ::| ::|
938 yLeuLysGlnAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 954
2051 AAGATTGAGAGCT 2062
      ||| ::| ::| ::| ::| ::| ::| ::| ::|
955 lYsLeuAlaAla 958

seq_name: p1r2:A82615
seq_documentation_block:
surface protein XP1981 [imported] - Xylella fastidiosa (strain 9a5c)
C/Species: Xylella fastidiosa
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C/Accession: A82615
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A/Reference number: A82515; MUID:20365717
A/Note: For a complete list of authors see reference number A59328 below
A/Accession: A82615
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1190 <SIM>
A/Cross-references: GB:AE004017; GB:AE003849; NID:9107083; PIDN:AAF84783.1; GSPDB:GN
A/Experimental source: strain 9a5c
R/Simpson, A.J.G.; Reinach, F.C.; Artuda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre
as-Melo, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A/Authors: Ferreira, V.C.A.; Perro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kempner, E.L.; Kitejima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins
A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
Rodrigues, V.; Rosa, A.J. de M.; de Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri,
A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tshako, M.H.; Valiada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A/Reference number: A59328
A/Contents: annotation
C/Genetics:
A/Gene: XF1981

alignment_scores:
Quality: 168.50 Length: 847
Ratio: 0.428 Gaps: 40
Percent Similarity: 46.517 Percent Identity: 21.842

alignment_block:
us-09-306-689-12 x A82615 ..
Align seg 1/1 to: A82615 from: 1 to: 1190
40 GCGCTGCGCGCTGCGAGCGGTTCTCAAGATTGGAGCTACGCGCTCGCTCC 89
      ||| ::| ::| ::| ::| ::| ::| ::| ::|
155 GlyLeuValAspLysSerGlyAla.....ThrYSerMetArgLe 168
90 GGGT.....GCGTCTAGCCAGCATTTGGAGCTACGCGCTCG 121
      ||| ::| ::| ::| ::| ::| ::| ::| ::|
168 uGlySerIleAlaThrMetAsnGlySerAlaGlyLysAspSerIleAlaThr 185
122 TCGCGCTCGCGAGCGGTAGCCAA.....GATTGAGCATTCGCGCTCG 162
      ::| ::| ::| ::| ::| ::| ::| ::|
185 lGlySerGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 201

```



```

      756 SerGly.....SerAsnAspValLeuPheAsnLeuSerSerSerLeu 770
      1623 ATG.....:|||||:|||||:|||||:|||||:|||||:|||||:
      770 sLeuAspLysLeuThrValGlyAspThrValMetThrThrAsnGlyVal 787
      1627 .....CTCAACAT 1635
      787 hrValGlySerGlyValThrLeuGlySerMetGlyLeuValIleThrAsp 803
      1636 AACAAATATGTGACA.....AACAGCTTAAGATTAATCTC 1673
      804 GlyProSerValThrSerSerGlyIleAsnAlaGlySerGlnLysIle 820
      1674 ATCTGTAAAGTCATTACCTGCTTAAGATTCGAGAAATGATTAGTCG 1723
      820 rAsnValAlaAlaGlyThrAlaAspThrAspAlaValAsnLeuSerGln 837
      1724 CTCCAACTTCATGTTGGATTCAAAGTTATCTTCTCTCAATTGCTAG 1773
      837 euAsnThrAlaMetAlaGlySerGlyAlaLysSerValHisTyr..... 851
      1774 GGATCTCAGCATTTGAGCTACGCGCTGCGCCTGGCAGCGGTTCTCAGA 1823
      852 ..TyrSerThrTyrAspGlyGlyThrGlnGlyGlyAsnTyrAsnGlyAs 867
      1824 TTGAGACTACGCGCTGCGCTGCGGCTGCTACGACGATTGAGCTAG 1873
      867 p.GlyAlaThrGlyThrArgSerIleAlaVal..GlyValGlyThrLeu 882
      1874 GCCTGCGCGCTGCGCAGCGGTACCAAGATTGAGCTACGCGCTGCGTCG 1923
      883 AlaSerAla.....GluGlyAlaThrAla..... 890
      1924 GGTGATCTCAGCATTTGAGCTACGCGCTGCGCCTGGCAGCGGTTCTCA 1973
      891 .....ValGlySerGlyAlaAlaAlaAlaSerGlyLys..... 900
      1974 AGATTGAGCTACGCGCTGCGCTGCGGCTGCTACGACGATTGAGCT 2023
      901 .....GlySerThrAlaIleGlyArgAsnAlaValAlaSerAlaAspGly 915
      2024 ACGGCTGCGCGCTGCGCAGCGGTACCAAGATTGAGCT 2062
      916 SerVal...AlaLeuGlyAspGlyAlaLysAspGlyAla 927

seq_name: p1r2:T14961

seq_documentation_block:
hypothetical protein Y1044 - Yersinia pestis plasmid pMT1
C:Species: Yersinia pestis
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 22-Oct-1999
C:Accession: T14961; T14657
R:Rindler, L.E.; Plano, G.V.; Burland, V.; Mayhew, G.F.; Blattner, F.R.
Infect. Immun. 66, 5731-5742, 1998
A:Title: Complete DNA sequence and detailed analysis of the Yersinia pestis KIM5 plasmid
A:Reference number: Z18268; MUID:99043898
A:Accession: T14961
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1525 <IN>
A:Cross-references: EMBL:AF074611; NID:g33883003; PID:g33883044; PIDN:AAC82704.1
R:Hu, P.; Elliott, J.; McCreedy, P.; Skovronski, E.; Garnes, J.; Kobayashi, A.; Cariano,
submitted to the EMBL Data Library, March 1998
A:Description: Structural organization of virulence determinants in three Yersinia pestis
A:Reference number: Z18168
A:Accession: T14657
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1525 <HUP>
A:Cross-references: EMBL:AF053947; NID:g2996286; PID:g2996290; PIDN:AAC13170.1
C:Genetics:

```

```

A:Gene: Y1044
A:Genome: plasmid pMT1

alignment_scores:
  Quality: 166.50      Length: 883
  Ratio: 0.386         Gaps: 44
  Percent Similarity: 48.811      Percent Identity: 21.065

alignment_block:
us-09-306-689-12 x T14961 ..

Align seg 1/1 to: T14961 from: 1 to: 1525

23 CTCAGCATTTGAGCTACGCGCTGCGCCTGGCAGCGGTTCTCAAGATTGG 72
|||||:|||||:|||||:|||||:|||||:|||||:
513 lIeSerLeuGlyLysValAlaAlaMetValAlaPhe..... 525
73 AGCTACGCGCTGCGGCTGCTACCGACGATTGAGCTACGCGC 121
526 .....LysLeuMetArgAlaGlyIleAlaGlyValIleGlyThrAlaG 540
122 .....TGCGCCCTGGCAGCGGTA 139
540 lylGlnMetValAsnThrPheThrLysMetSerThrValLeuGlnAlaPro 556
140 GCCAAGATTGAGCTACGCGCTGCGCTGCGGCTGAGATTCAGCATT... 184
557 PheAsnLeuGlyAlaThrAlaValThrArgPheAsnArgAlaAlaArgme 573
184 ..... 184
573 tGlyLeuAlaProIleProSerLeuIlePheAlaIleArgGlyAlaIleT 590
185 .....GGAGCTACGCGC.....TGCGCCCTGGCAGCGGTTCTC 217
590 hnglyLeuGlnGlyAlaPheAlaGlyLeuThrAlaPheIleAlaAsn 606
218 AAGATTGAGCTACGCGCTGCGTC..CGGCTGCTTACGCCAGCATT... 262
607 ProIleGlyAlaAlaPheThrValAlaThrValAlaValAlaGlyLeuI 623
263 .....GGAGCTACGCGCTGCGCTGCGCAGCGG 290
623 eThrTyrMetThrMetLeuArgSerGlyThrSerLysValValAspGlu 640
291 TAGCCAGATT.....GGAGCTACGCGCTGCGCTGCGGCTGATTCAGCT 334
640 lArgGlyIleProGlnAlaMetThrAlaAlaLysArgAlaGlnMetAla 656
335 TCCCA..... 339
657 AlaArgAlaAlaGlyLeuGlnLysGlnIleGlnArgAspGlnAlaIle 673
340 .AAACTGGGCAAAATAATATCTCTATATCTCCCAATATTACAT 388
673 uylStrnGlyGluSer.....ValAsnTyr...T 682
389 ATGATCTGTAACAAGCTAATGTTTACAGATTATGCAAGCGCGCA 438
682 ySerThrAlaAlaGly.....ProValAlaValLysGluSerLys 695
439 GAGTTGGGATTTGAGGTACAAAGAGAAGACCAATATATTGCAACAC 488
696 GluVal.....ValGlnAlaArgLeuLysLysAsnGlnGluLys 709
489 TCAAACCACTTACGACGATTCAAACCGCTATTTGCTTACACGACG 538
709 rGluArgThrThrGlyThrMetAlaLeuGlyAspGlyAlaValAlaLys 726
539 GCATTGTGTTATCCGCTCCCAATATGATTAATTTGCTACAGAACTAA 588
726 rGluAlaLysGluAla.....AlaGluSerGlnIleGlnLysIleArg 740

```


A:Title: A lipopolysaccharide-binding domain of the Campylobacter fetus S-layer protein
 A:Reference number: A56143; MUID:95204338
 A:Accession: A56143
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1109 <DMC>
 A:Cross-references: GB:S57860; NID:913763; PIDN:AAB3871.1; PID:913764
 A:Note: this gene appeared to be silent in strain 82-40 LP3
 C:Genetics:
 A:Gene: sapA2

alignment_scores:
 Quality: 165.00 Length: 672
 Ratio: 0.493 Gaps: 34
 Percent Similarity: 49.851 Percent Identity: 22.917

alignment_block:
 US-09-306-689-12 x A56143 ..

Align seg 1/1 to: A56143 from: 1 to: 1109

```

346 GGGGCAAAAAAATTATCTCTATATTTCCCAAAATTACCAATATGATAC 395
    ||| |||
    ||| |||
23 GlysnlystyrtpleuasptryAlaasnasnserleuglyValse 39
396 TGAACAAGTAAATGCTTACAGATTAGTCAAGCGCCGCAAGATTGG 445
    ||| |||
    ||| |||
39 rserleualasnilemetleuaspserproglyAlaAlaLysphephec 56
446 GGATTGAGTACAAAGAGAACGCAATATATTCAGAACGCTCAAAACC 495
    ||| |||
    ||| |||
56 LysaspserleuAlaGlyasnleulysasphevalthrLysiletyr 72
496 AGTTAGGACAGATTCAACCGCTATTGGCTTAACGAGCGTGCATTTG 545
    ||| |||
    ||| |||
73 Serlle.....AlaleuglyAsnthrser..... 80
546 GTTATCCGCTCCACAAATGCTAATGCTACAGAAACCTRAAGCA.... 591
    ||| |||
    ||| |||
81 .....AspyalaspglylleasntryttrpThrLysAlaIleT 93
592 .....GGCCAAGCATTTAGTTGCTCCGCAAGCATTTGTACAAATGCA 633
    ||| |||
    ||| |||
93 hreglyglygluIuphetraspserlysglyasnvalIleserValAla 109
634 AATTAACCAAACTGTATTATCGCATTCATCAATTTTATGGCTCAGT 683
    ||| |||
    ||| |||
110 SerleuaserlysglyAspleuilegly.....Alametlleasnserme 124
684 ATTGGCTGGAAATGATTTACAT.....GAGGCTTTACAGA 718
    ||| |||
    ||| |||
124 tvalasnlyglyserAlaIleuaserlyAlaIlephegluAlaLysAla 141
719 ATTAACAGCAACCAACATGCTCTGCTTAAGCTGGCTGGAGCT.... 762
    ||| |||
    ||| |||
141 lalalaserasptyrpheAlaaspAlaIleuIleuglyLysaspIleSerGly 157
763 .....ACAAATTCATTAATTAATAATTTGCTAATTCAGT 797
    ||| |||
    ||| |||
158 leuaspIleuglythrthrserlyserlyserleuileSerGluIle...asnserAl 173
798 AAAAACCTTGACGAATTTGGTGAAGCAATTAAGTCAATTTGGTTCAAAAC 847
    ||| |||
    ||| |||
173 aserAspIleuaspLysValLysSerGluIle..... 183
848 TACAAATATCAAAAGCTTAGGACTTAGAGACAAACCTCAAAATATC 897
    ||| |||
    ||| |||
184 .....AspglyleuLysGlu... 188
898 GGTGGACTGTATAAGCTGGCCTT..... 921
    ||| |||
    ||| |||
189 .....SerlleaspIleuIleuglyleuAsnlyslleAlaIleuThrthrGluAs 204
```

```

922 .....GGTTAGATGTATATCTACAGCGCTATAT 949
204 naspthrIlethrlythrlythrlygllyAspleuileSerlelyValG 221
950 CGGGCGCAACAGCTGCATCTGTACTGTACAGATTAATAATGCTCAACAGCT 999
    ||| |||
    ||| |||
221 lythrAla.....AlaIleuaserlythrleuAsnprogly 231
1000 AAAAAGTG.....GGTGGGGT.....TTGAATTGGC 1028
    ||| |||
    ||| |||
232 AspylleaspIleuglyAlaIleuAsnaspValleuLysValaspLeuLy 248
1029 AAACCAAGTTGTGCT.....AATATTACCA 1054
    ||| |||
    ||| |||
248 sasnasnphelysglyleuLysaspaspglytyrIleLysasnIleGluL 265
1055 AA.....GCCGTTCTCTTACATTTAGCCCAACGT 1086
    ||| |||
    ||| |||
265 yslleuSerleuThrAsnSerSerValSerasnAlythrPheaspAlaLys 281
1087 GTTGCACAGGTTTATCTCAACTGGCGCTGCGCTCTTAATTTGCTTC 1136
    ||| |||
    ||| |||
282 glylleaspIleuglyInthr.....ValAlaIle 291
1137 TACTGTTCTCTCCGATTAGCCCATTTAGCTTGGCGTATTCGCAT 1185
    ||| |||
    ||| |||
291 userglyleuLysglylleSerValthrAsnleuAlaasnIleValaspv 308
1186 ..AATTTAATCATGCAAA.....AGTTAGAGAGT...TAT 1218
    ||| |||
    ||| |||
308 algluValasnGlyPheylsglyThrAsnphasnValaspSerIletyr 324
1219 GCCGACGCTTTAAAAAATTAGCTATGACGAGCAATATTATTATTCAGCA 1268
    ||| |||
    ||| |||
325 AlaaspysValleuAsnaspIlyserAlaspValalGlnasnleuLysVal 340
1269 ATATACACGGGGCAACAGGACTATTGATCGATCGCTACT..... 1308
    ||| |||
    ||| |||
341 .....AsnGlyValalGlyAlaLysGlyAlaSerValAlaIlethrAla 355
1309 ..GCAATTAATACCGCATTTGGCGCTATGCTGGTGGTGGTGGCT 1356
    ||| |||
    ||| |||
355 splysIleGluIthrLeuasnleuAsnThrThrlySerGlnserPheVal 371
1357 GCAGCGATTTTACATTTGAAAGTAAACATTAATCTTGCATCACGAA 1406
    ||| |||
    ||| |||
372 SerAlaspValAlaSerIleSerVallysglyAsnAlaasnleuSerle 388
1407 TAGCAAAAAAGACAAAGTGCACATTCAAAACCTGGTTCCGAGAGGCTGATT 1456
    ||| |||
    ||| |||
388 uAlaIthrGlyAlaIlythrThrleuAsp.....AlaserSerp 402
1457 TTGCTAAGAAGGCTTAATTAATAAGCACTAAGATGAGAAATGCA 1506
    ||| |||
    ||| |||
402 heglyglyAlaIleuaspAlaspLeuSerThr...SerAlaSerValThr 417
1507 GAATCATCGCTCAAAATGCG...GAGCGGATCACCTCAACGAAGTT... 1551
    ||| |||
    ||| |||
418 SerllelysglyAsnGlyAsnaspIlysllethrIleLysaspValAl 434
1552 .....GATGACTTATGCA 1567
    ||| |||
    ||| |||
434 aValasnValAlaIleaspIlyglyAlaIleuAsnaspIleuValIleL 451
1568 AAGGTACGGCAAAATTAACCAAGATGAGTATCAAAAGTTGTGATAC 1617
    ||| |||
    ||| |||
451 ysglySer.....ThrAlaspThrleuGlnProthrleuThrAsn 464
1618 TATGAA.....TTGCT 1628
    ||| |||
    ||| |||
465 IlegluLysValthrIleaspGlyAsnThrLysaspLeuThrleuSerle 481
```



```

301 PheLeu.....ValGlyGlyAlaLeuThrThrAsnAspProthrGrl 314
1122 TGCCTTAATGCTTCTACTGTTCTCTT..... 1149
314 yserThrProLathrMetSerSerLeuPheYsAlaAlaAspAspLysA 331
1150 .....GCGATTAGCCCATTCAGATTGGCCGGTATTCCGCTAAATTAA 1194
331 sPAlaAlaLeuInsSerTLeasPheGlyGlyLysLysTyrGluPheAla 347
1195 CATGCAAAAAGTTTAGAGATTATGCCGACGCTTAAAAAATTAGCTA 1244
348 GLyGlyAsnSerThrAsnGlyGlyValLysPheLysAspThr...Va 363
1245 TCACGGAGATTAATTATTCAGCAATATACGCGGAGACAGGACTATTG 1294
363 lSerSerAspAlaLeuLeuAlaGlnValLysAlaAspSerThrAlaAsnA 380
1295 ATGCATCGGTTACTGCAATTAAATACCGCATG.....GCCGCTATT 1335
380 snValLysLleThrPheAsnAsnGlyProLeuSerPheThrAlaSerPhe 396
1336 GCTGGTGATGCTCTGCTGCTGCTCAGCCGATTAAACATTGAAAAAGTTAA 1385
397 GlnAsnGlyValSerGlySerAlaAlaSer..... 406
1386 ACATAATCTTGTCATCAGCAATAGCAAAAAAGAGAAAGTACCATTCAAA 1435
407 ...AsnAlaAlaTyrLleAspSerGluGlyLeuThrThrThrGluS 422
1436 ACTGGTCCGAGAGGCTGATTTTGGTAAAGAACTGCTTAATTATAAGCA 1485
422 eTyr.....AsnThrAsnTyrSerVal 429
1486 ACTAAAGATGAGAAAAATCGAAGAAATCATCGTCAAAATGCGAGCGAGAT 1535
430 AspLysAspThrGlyAlaValSerValThrGlyGlySerGly..... 443
1536 CACCTCAAAAGCAAGTTGATGATCTTATGCCAAA.....GGTA 1573
444 ThrGlyLysTyrAlaAlaAsnValGlyAlaGlnAlaTyrValGlyAlaA 460
1574 ACGGCAAAATTCACCAAGATGAGCTATCAAAAGTCTTATACATTCGAA 1623
460 sPGLysLysLeuThrThrAsnThrThrSerThrGly..... 471
1624 TTGCTCAACATAGCAAAAATGTGACAAACAGCTTAGATTAAGTATCTC 1673
472 .....SerAlaThrLysAspProLeuAsnAlaLeuAspGlnAlaLeuAl 486
1674 ATCTGTAAAGTCATTACCTCGTCTAATGATTTCGAGAAATGATTATAGTG 1723
486 aserTLeasPlysPheArgSerSer.....LeuGlyA 497
1724 CTCCAACTTCATGTTGGATCAAAAGTTTATCTCTCTT 1761
497 lAlleGlnAsnArgLeuAspSerAlaValThrAsnLeu 509
seq_name: p1r2:I37984
seq_documentation_block:
keratin 9, type I, cytoskeletal - human
N:Alternate names: Cytokeratin 9; scatter protein 60k chain, placental
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence.revision 02-Jul-1996 #text.change 29-Oct-1999
C:Accession: I37984; S40307; S41161; B35494; I37943
R:Reis, A.; Hennies, H.C.; Langbein, L.; Digweed, M.; Mischke, D.; Drechsler, M.; Schiroc
Nature Genet. 6, 174-179, 1994
A:Title: Keratin 9 gene mutations in epidermolytic palmoplantar keratoderma (EPPK).
A:Reference number: I37984; MUID:94214498
A:Accession: I37984
A:Status: preliminary; translated from GB/EMBL/DBJ

```

```

A:Molecule type: DNA
A:Residues: 1-622 <RES>
A:Cross-references: EMBL:X75015; NID:9453154; PIDN:CAAS2924.1; PID:9453155
R:Langbein, L.; Heid, H.W.; Moill, I.; Franke, W.W.
Differentiation 55, 57-71, 1993
A:Title: Molecular characterization of the body site-specific human epidermal cytoker
A:Reference number: I37943; MUID:94131202
A:Accession: S40307
A:Molecule type: mRNA
A:Residues: 1-622 <LAN>
A:Cross-references: EMBL:Z29074; NID:9435475; PIDN:CAA82315.1; PID:9435476
A:Accession: S77921
A:Molecule type: protein
A:Residues: 29-53;62-104;167-188;199-233;241-249;295-362;374-430;450-480;579-604 <LAF
R:Langbein, L.
submitted to the EMBL Data Library, December 1993
A:Reference number: S41161
A:Accession: S41161
A:Molecule type: mRNA
A:Residues: 1-11,'SR',13-622 <LAN>
A:Cross-references: EMBL:Z29074; NID:9435475; PIDN:CAA82315.1; PID:9435476
R:Rosen, E.M.; Meromsky, L.; Romero, R.; Selter, E.; Goldberg, I.
Biochem. Biophys. Res. Commun. 168, 1082-1088, 1990
A:Title: Human placenta contains an epithelial scatter protein.
A:Reference number: A35494; MUID:90267446
A:Accession: B35494
A:Molecule type: protein
A:Residues: 'X',450-465 <ROS>
C:Genetics:
A:Gene: GDB:KRT9; EPPK
A:Cross-references: GDB:303970; OMIM:144200
A:Map position: 17q12-17q21
A:Introns: 213/3; 241/2; 293/3; 347/3; 389/3; 464/2
A>Note: defects in this gene may cause epidermolytic palmoplantar keratoderma
C:Superfamily: cytoskeletal keratin
C:Keywords: coiled coil; intermediate filament
F:1-153/Domain: head #status predicted <HEA>
F:154-459/Domain: helical rod #status predicted <ROD>
F:460-622/Domain: tail #status predicted <TAI>

alignment_scores:
Quality: 163 00 Length: 741
Ratio: 0.487 Gaps: 33
Percent Similarity: 45.209 Percent Identity: 20.648

alignment_block:
US-09-306-689-12 x I37984
Align seg 1/1 to: I37984 from: 1 to: 622

22 TCTCAGCATTTGAGCTACGGCTCGCCCTGCGACGCT.....TCTCA 65
|||||: ||| ||| ||| |||: |||:
32 SerArgPheSerSerSerGlyArgGlyGlyGlyArgPheSerSe 48
66 ACATTGGAGCTACGGCTCGCTCCGGGTGCTTACCCAGCATTTGGAGCT 115
|||||: ||| ||| ||| |||: |||:
48 rSerSerGlyTyrGly.....GlyGlySerSerArgValGlyGlyA 62
116 ACGGCTGGCCCTGCGACGCTAGCAAGATTGGAGTACGCGCTGGCT 165
|||||: ||| ||| ||| |||: |||:
62 rGly.....GlyGlyGlySerPheGlyTyrSerTyrGlyGly 75
166 CCGGCTGATCT...CAGCATTTGAGCTACGGCTCGCCCTGCGACGCG 212
|||||: ||| ||| ||| |||: |||:
76 SerGlyGlyGlyPheSerAlaSerSerLeuGlyGlyGlyPheGlyGly 92
|||||: ||| ||| ||| |||: |||:
213 TTCTCAAGTTGGAGCTACGGCTCGCTCCGGGTGCTTACCCAGCAT 262
|||||: ||| ||| ||| |||: |||:
92 YserArgGlyPheGly...GlyAlaSerGlyGlyGlyTyrSerSerSe 108
GGAGCTACGCGCTGCGCCCTGCGACGCTAGCAAGATTGGAGTACGCGC 312
|||||: ||| ||| ||| |||: |||:

```

```
108 LysIlyPheglyglyIlyPheglyglyIlySer.....GlyglyIly 121
313 CTGCTGCTCCGGT...GGATCTAGCTTCCCAAAACCTGGGCAAAAAAAT 359
    :::: ||| |||||::: ||:::
122 PheglyglyIlyTyrglySerglyPheglyglyLeuSlygly..... 135
360 TATCTCTATATTCGCCAAATACCAATATGATCTGAACAAGTAATG 409
136 .....PheglyglyIlyAlaIlygly 143
410 GTTACAGATTAGTCAAAAGCCGCAAGATGGGATGAGTCAAA 459
    ||::: ||| ||:::
143 LyspIlyglyIlyLeuThrAlaasnIlyS.....SerrThmet 156
460 AGAAGAAGCAATATATTGCAACAGCTCAAAACGTTAGGCAGAT 509
    :::: ||| ::::: ||:::
157 GluIlyLeuasnSerArgLeuAlaSer..... 165
510 TCAACCGCTATTGGCTTAACGCTGCGCATTTGTTATCCGCTCCAC 559
166 .....T 166
560 AAATTGATAATTGCTACAGAAACTTAACAGCCAGCATTAAGTTCT 609
    :::: |||||::: |||||:::
166 TyLeuaspIlyVal.....GlnAlaLeuGlu 175
610 GCCGAAAGCATTTGACAAATGCAATTAAGCCAAACTGTATTATCTGG 659
    ||::: ||:::
176 Alaasn..... 178
660 CATTCATCTATTTAGCTCAGTATTGGCTGGAATGATTAGATGAG 709
    |||||:::
179 .....AspLeuGluasnL 183
710 CCTACAGAAATACAGCAACCAACATGCTCTGCTAAAGCTGGCTTGAG 759
183 yslIleGlnaspIlyTyraspIlySglyProAla..... 194
760 CTACAAATTCATTAAATGAAATATGTCAATTCAGTAAACACTTGA 809
    :::: |||||::: ||:::
195 .....AlaIleGlnIlyasnTyrrSerProTyrrIlyasnThrIleasn 208
810 CGAATTTGGTGCAAAATAGTCAATTTGGTCAAAACTACAAATATCA 859
    ||::: |||||::: ||:::
208 pasPleuIlyAspGlnIlyValaspLeu..... 217
860 AAGCTTAGGACTTAGAGACAACTCAAAATATCGGTGAGCTTGAT 909
    ||::: |||||::: ||:::
218 .....ThrValGlyIlyasn.....LysThrLeuLeuaspIlyLeasp 229
910 AAGCTGCTGCTTGTAGATTATCTCAGGCTATTATCGGGCGCAAC 959
    :::: |||||:::
230 AsnThrIlyMetThrIleuaspPheArg..... 239
960 AGCTGCACTTGTACTTCAGATAAAATGCTCAACAGCTAAAAAATG 1009
    :::: ||::: ||:::
240 .....IleLysPheGluMetGluIlyAsnLeuArgGlnIlyVala 253
1010 GTGCGGCTTTGAATGGCAAAACCAAGTTGGTAAATATTACCAAGCC 1059
    |||::: |||||::: ||:::
253 spAlaaspIlyLeasnGlyLeuArgIlyValleuaspIlyLeuThr..... 267
1060 GTTCTCTTACATTTAGCCCAACGTTGCAAGCTTATCTTCAAC 1109
267 ..... 267
1110 TGGGCTGTGCTGCTTAAATGCTTCTACTGTTCTTGCGATTAGCC 1159
    :::: ||| ::::: ||:::
268 .....MetGluIlySerraspLeuGluMetGlnIlyclur 279
1160-CATTAGCATTTGCCGGTATTGCCGATAATTTAATCAGCAAAAAGTTTA 1209
    |||::: |||||::: |||||:::
279 hrLeuGlnIlyGluLeuMetAlaLeuIlySlyasnIlySglyGluMet 295
```

```
1210 GAGAGTATCCGACGCTTTAAAAATATGAGCTATGACGAGATAATT 1259
    :::: ||| |||||::: |||||:::
296 SerGlnLeuThr.....GlyGlnasn.. 302
1260 ATTAGCAGAAATATACGGGGAAACAGGACTAATGATGATGCGTTACTG 1309
    :::: |||||::: |||||:::
303 .....SerrGlyaspValaIlyValIlyLeasnV 312
1310 CAATTATACCGCATTCGCGCTATTGCTGTGCTGTCTGCTGCCA 1359
    :::: ||| ::::
312 al.....AlaProGly 315
1360 GCCGATTTAATCATTTGAAAAAGTTAAACATAATCTTGTCTACGATAAG 1409
    |||||::: |||||:::
316 LysaspLeuThr.....LysThrIleuasnAspMet.....ArgI 327
1410 CAAAAAAGAGAAAGTACCATTCGTTCCGAGAGCGTGAATTTTG 1459
    :::: |||||::: |||||:::
327 ngIlyTyrgIlyGlnLeuIlyAlaIlyasn.....A 337
1460 CTAAAGAGTGCCTAATTATTAAGCAACTAAAGATGAGAAATCGAAGAA 1509
    |||||::: |||||:::
337 rglYaspIlyGlnaspGlnIlyTyrgIlyThrGlnIlyGlnIlyHis 353
1510 ATCATCGCTCAAAATGCGAGCGATCACCTCA.....AAGCAAGTGA 1553
    :::: |||||::: |||||:::
354 GluValSerSerSergIlyGlnIlyValGlnSerSerAlaIlySgluValIth 370
1554 TGATCTTATGCAAAAGTACGCGCAAAAT.....ACCCAAGATG 1594
    :::: ||| ::::
370 rglIlyLeu...ArgHisGlyValGlnIlyGlnIlyGlnIlyGlnIlySerg 386
1595 AGCTATCAAAAGTTGTGATAACTGATGCTCAACATACGCAAAAT 1644
    :::: |||||::: |||||:::
386 IlyLeuSerIlySlyAlaAlaLeuGlnIlySerLeuGlnaspIlyIlyasn 402
1645 .....GTGACAAACAGCTTAGATTAAGTATCATCTGTAAG 1682
403 ArgTyrrCysGlyGlnLeuGlnMetIleGlnGlnIlyIleSerrsnLeuG 419
1683 TGCA.....TTTACCTGCTTAATGATTCGA 1708
    |||::: |||||:::
419 uAlaGlnIlyThrAspValArgGlnIlyGlnIlyCysGlnIlyGlnIly 426
1709 GAATGATATAGTGCCTCCAACTCAATGTTGATGATCAAAATTAATCTCT 1758
    :::: |||||::: |||||:::
436 yrrSerLeuLeuSerIleLysMetArgLeuGlnIlySgluIlyGluThr 452
1759 CTTCAA..... 1764
453 TyrHisasnLeuLeuGlnIlyGlnIlyGlnIlyaspPheGluSerSergIlyAl 469
1765 .....TTTGTAGGGGATCTCAGATTTGAGTACGCGCC 1798
    :::: |||||::: |||||:::
469 agIlyIlyIleGlyLeuGlnIlyArgIlySergIlySerrTyrrGly.. 485
1799 TGGGCTTGGCAGCGGTTCTCAAGATTGGAGTACGCGCTCGCGGCT 1848
    |||::: |||||::: |||||:::
486 ..ArgIlySerArgGlySergIlySerrTyrrGlySergIlySerrTyrr 501
1849 GCGCTTAGCCAGCATTTGAGCTACGCGCTCGCGCTGAGCGGTAAGCA 1898
    |||||::: |||||::: |||||:::
502 GlyIlyIlyTyrrIlyGlnIlySergIlySerrArgGlySerrIlyGly.. 517
1899 AGATTGAGCTACGCGCTCGCTCGCGGTGAGTCAAGATTTGAGCTACG 1948
    :::: ||| ::::: |||||:::
518 ..SerrTyrrGlyIlySergIlySerrGlySergIlySergIlyGlyIlyTyrr 534
1949 GCGTGGCGCTTGGCAGCGGTTCTCAAGATTGAGCTACGCGCTCGCGG 1998
    |||::: |||||::: |||||:::
534 IlyIlySergIlyGlyIlyHisSerrIlySerrIlySergIlySerrIlyHisSerr 550
```



```
1182 leProLeuLysMetPheAsnValAlaLeuThrMetMetThrProAlaLeu 1198
1225 CGCTTTAAAAATTGAGTATGACGAGATATTATTATAGCAGATATCA 1274
1199 MetMetGlyAlaAlaGlyPheValArgPheAsn..... 1209
1275 GCGGGGACAGGAGCTATTGATGATCGGTTACTGCAATTAAATACCGCAT 1324
1210 .....AlaGlnValMetGlnSerThrAlaGly 1219
1325 TGGCCGCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1374
1219 etSerThrLeuSerGlyAlaIleSerThrIleProAlaLysLeuThr... 1234
1375 GAAAAAGCTTAACATTAATCTTGTCATCAGCAATATGCAAAAAAGAGAAAGT 1424
1235 .....SerValAlaSerSerAlaAsnThr 1242
1425 GACCATTCAAAACCTGTTCCGAGAGGCTGATTGCTAAAGAAAGTGCTTA 1474
1242 rThrThrLeuSerIleMetArgIleAlaThrSerAlaProArgIleAla 1259
1475 ATTTAAAGCACTAAAGATGAGAAAATCGAAGAAATCATCGTCAAAAT 1524
1259 erAlaMetSerSerAlaAlaGlyGlnValGlnSerAlaMetGlnArgMet 1275
1525 GCGGAGCGGATCACCTCAAGCAAGTTGATGATTCGCAAAAGGTAA 1574
1276 AlaGlnSerValGlnSerSer..GlnTyrMetIleGlnMetGlyArg 1291
1575 CCGCAAAATTACCCAGATGAGCTATCAAAAGTTGATTAATATGAAAT 1624
1291 gGlnAlaGlyThrGlnThr..... 1297
1625 TGCTCAACATAGCAAAATGTGCAACAGCTTACATTAATATCTCA 1674
1298 .....GlyGlnAsnIleAlaAsnGlyIleLys.....Ser 1307
1675 TCTGTAGTGCATTTACCTCGTCAATGATTCGAGAAATGATTAAGTGGC 1724
1308 SerValGlyAlaValSerSerAlaValAspSer....LeuValAsnAl 1322
1725 TCCAACTTCATGTGGATCAAGTTATCTTCTTCAATTGCTAGGG 1774
1322 aaIaArgAlaArgAlaMetGlyGlyValGlyAlaMetGlnIleVal.... 1337
1775 GATCTGACGATTTGAGCTACGGCTGCGCCCTGCGACGGGTTCTCAAGAT 1824
1338 .....GlyAlaMetIleGlyLysGly..LeuAlaAs 1347
1825 TGGAGCTACGCGCTGCGTCCGGGTGCTAGCCAGCATGAGCTACGG 1874
1347 nGlyMetIleAla.....SerLeuGlyAlaValThrAlaAla 1360
1875 CCTGCGCCCTGGCAGCGGTAGCCAGATTTGAGCTACGGCC 1915
1360 laAsnAlaLeuValAlaGlnAlaGlnArgAlaAlaGlnAla 1373
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:54:36 ; Search time 47.48 Seconds
(Without alignments)
18,532 Million cell updates/sec

Title: US-09-306-689-11
Perfect score: 290
Sequence: 1 QHMSYGLRGSGSQDMSYGLR.....SYGLRPGSGSQDMSYGLRPG 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued Patents AA:*
2: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/PCYUS.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290	100.0	49	1 US-08-387-156-4	Sequence 4, Appli
2	290	100.0	49	2 US-08-694-865-4	Sequence 4, Appli
3	290	100.0	49	3 US-08-878-748-4	Sequence 4, Appli
4	290	100.0	49	4 US-09-124-491-4	Sequence 4, Appli
5	290	100.0	544	1 US-08-387-156-10	Sequence 10, Appli
6	290	100.0	544	2 US-08-694-865-10	Sequence 10, Appli
7	290	100.0	544	3 US-08-878-748-10	Sequence 10, Appli
8	290	100.0	544	3 US-09-124-491-10	Sequence 10, Appli
9	290	100.0	699	3 US-08-694-865-16	Sequence 16, Appli
10	290	100.0	699	3 US-09-124-491-16	Sequence 16, Appli
11	290	100.0	977	2 US-08-387-156-8	Sequence 8, Appli
12	290	100.0	977	2 US-08-694-865-8	Sequence 8, Appli
13	290	100.0	977	2 US-08-878-748-8	Sequence 8, Appli
14	290	100.0	977	3 US-09-124-491-8	Sequence 8, Appli
15	187.5	64.7	44	1 US-07-690-983D-45	Sequence 45, Appli
16	187.5	64.7	84	1 US-07-690-983D-47	Sequence 47, Appli
17	152	52.4	55	3 US-08-458-814-7	Sequence 7, Appli
18	150.5	51.9	52	3 US-08-458-814-6	Sequence 6, Appli
19	104.5	36.0	24	1 US-07-690-983D-43	Sequence 43, Appli
20	96	33.1	20	1 US-07-690-983D-40	Sequence 40, Appli
21	74	25.5	90	1 US-08-341-219-19	Sequence 19, Appli
22	72	24.8	18	1 US-07-690-983D-20	Sequence 20, Appli
23	70	24.1	16	1 US-08-453-588-26	Sequence 26, Appli
24	70	24.1	16	3 US-08-521-079-26	Sequence 26, Appli
25	69	23.8	17	1 US-07-690-983D-16	Sequence 16, Appli
26	67	23.1	14	1 US-08-453-588-23	Sequence 23, Appli
27	67	23.1	14	3 US-08-521-079-23	Sequence 23, Appli
28	67	23.1	936	1 US-08-455-970A-12	Sequence 12, Appli

29	66.5	22.9	89	1	US-08-341-219-20	Sequence 20, Appli
30	66	22.8	14	1	US-07-690-983D-24	Sequence 24, Appli
31	66	22.8	18	1	US-07-690-983D-28	Sequence 28, Appli
32	66	22.8	26	2	US-08-446-692-29	Sequence 29, Appli
33	66	22.8	26	2	US-08-488-351A-29	Sequence 29, Appli
34	66	22.8	750	5	5457037-3	Patent No. 5457037
35	66	22.8	751	5	5457037-5	Patent No. 5457037
36	65.5	22.6	17	1	US-07-690-983D-18	Sequence 18, Appli
37	65	22.4	17	1	US-08-188-223-5	Sequence 5, Appli
38	65	22.4	17	3	US-08-968-466-5	Sequence 5, Appli
39	65	22.4	29	1	US-08-446-692-28	Sequence 28, Appli
40	65	22.4	29	2	US-08-488-351A-28	Sequence 28, Appli
41	65	22.4	34	1	US-08-446-692-16	Sequence 16, Appli
42	65	22.4	34	2	US-08-488-351A-16	Sequence 16, Appli
43	65	22.4	37	2	US-08-446-692-24	Sequence 24, Appli
44	65	22.4	37	2	US-08-488-351A-24	Sequence 24, Appli
45	64	22.1	12	1	US-08-453-588-24	Sequence 24, Appli

ALIGNMENTS

RESULT 1
US-08-387-156-4
Sequence 4, Application US/08387156
Patent No. 5723129
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HOW P. A.
TITLE OF INVENTION: GRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/387,156
FILING DATE: 10-FEB-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-387-156-4
Query Match 100.0%; Score 290; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 5.3e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491-4

Query Match 100.0%; Score 290; DB 3; Length 49;
Best Local Similarity 100.0%; Pred. No. 5.3e-27;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 OHWSYGLRPGSSQDWSYGLRPGSSQHSYGLRPGSSQDWSYGLRPG 49
Db 1 OHWSYGLRPGSSQDWSYGLRPGSSQHSYGLRPGSSQDWSYGLRPG 49

RESULT 5

US-08-387-156-10
Sequence 10, Application US/08387156
Patent No. 5723129

GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.

APPLICANT: REDMOND, MARK J.

APPLICANT: HUGHES, HOW P. A.

TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: REED & ROBINS

STREET: 635 BRYANT STREET

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: UNITED STATES OF AMERICA

ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/387,156

FILING DATE: 10-FEB-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/960,932

FILING DATE: 14-OCT-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/779,171

FILING DATE: 16-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: ROBINS, ROBERTA L.

REGISTRATION NUMBER: 33,208

REFERENCE/DOCKET NUMBER: 9001-0016.21

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 617-8999

TELEFAX: (415) 327-3231

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 544 amino acids

TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-387-156-10

Query Match 100.0%; Score 290; DB 1; Length 544;
Best Local Similarity 100.0%; Pred. No. 8.1e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 OHWSYGLRPGSSQDWSYGLRPGSSQHSYGLRPGSSQDWSYGLRPG 49
Db 494 OHWSYGLRPGSSQDWSYGLRPGSSQHSYGLRPGSSQDWSYGLRPG 542

RESULT 6

US-08-694-865-10
Sequence 10, Application US/08694865
Patent No. 5837268

GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.

APPLICANT: MANN, JOHN G.

TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS

NUMBER OF SEQUENCES: 34

CORRESPONDENCE ADDRESS:

ADDRESSEE: REED & ROBINS LLP

STREET: 285 HAMILTON AVENUE, SUITE 200.

CITY: PALO ALTO

STATE: CA

COUNTRY: USA

ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/694,865

FILING DATE: 09-AUG-1996

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: MCCracken, THOMAS P.

REGISTRATION NUMBER: 38,548

REFERENCE/DOCKET NUMBER: 9001-0016.22

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415)327-3400

TELEFAX: (415)327-3231

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 544 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

Query Match 100.0%; Score 290; DB 2; Length 544;
Best Local Similarity 100.0%; Pred. No. 8.1e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 OHWSYGLRPGSSQDWSYGLRPGSSQHSYGLRPGSSQDWSYGLRPG 49
Db 494 OHWSYGLRPGSSQDWSYGLRPGSSQHSYGLRPGSSQDWSYGLRPG 542

RESULT 7

US-08-878-748-10
Sequence 10, Application US/08878748
Patent No. 5969126

GENERAL INFORMATION:

APPLICANT: POTTER, ANDREW A.

APPLICANT: REDMOND, MARK J.

APPLICANT: HUGHES, HOW P. A.

TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS

NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,748
FILING DATE: 19-JUN-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016, 21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-748-10

Query Match 100.0%; Score 290; DB 2; Length 544;
Best Local Similarity 100.0%; Pred. No. 8,1e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSDMSYGLRPGSSOHMSYGLRPGSGSDMSYGLRPG 49
DB 494 QHWSYGLRPGSGSDMSYGLRPGSSOHMSYGLRPGSGSDMSYGLRPG 542

RESULT 8
US-09-124-491-10
Sequence 10, Application US/09124491
Patent No. 6022960
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANNIS, JOHN G.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,491

FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016, 22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491-10

Query Match 100.0%; Score 290; DB 3; Length 544;
Best Local Similarity 100.0%; Pred. No. 8,1e-26;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSDMSYGLRPGSSOHMSYGLRPGSGSDMSYGLRPG 49
DB 494 QHWSYGLRPGSGSDMSYGLRPGSSOHMSYGLRPGSGSDMSYGLRPG 542

RESULT 9
US-08-694-865-16
Sequence 16, Application US/08694865
Patent No. 5837268
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANNIS, JOHN G.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,865
FILING DATE: 09-AUG-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016, 22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 699 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-865-16

Query Match 100.0%; Score 290; DB 2; Length 699;
Best Local Similarity 100.0%; Pred. No. 1,1e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OHMSYGLRPGSGSDMSYGLRPGSSQHMYSYGLRPGSSQDMSYGLRPG 49
|||||
DB 9 OHMSYGLRPGSGSDMSYGLRPGSSQHMYSYGLRPGSSQDMSYGLRPG 57
|||||

RESULT 10
US-09-124-491-16
Sequence 16, Application US/09124491
Patent No. 6022960
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,491
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 699 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491-16

Query Match 100.0%; Score 290; DB 3; Length 699;
Best Local Similarity 100.0%; Pred. No. 1,1e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 OHMSYGLRPGSGSDMSYGLRPGSSQHMYSYGLRPGSSQDMSYGLRPG 49
|||||

DB 9 OHMSYGLRPGSGSDMSYGLRPGSSQHMYSYGLRPGSSQDMSYGLRPG 57

RESULT 11
US-08-387-156-8
Sequence 8, Application US/08387156
Patent No. 5723129
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HOW P. A.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/387,156
FILING DATE: 10-FEB-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.
REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 977 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-387-156-8

Query Match 100.0%; Score 290; DB 1; Length 977;
Best Local Similarity 100.0%; Pred. No. 1,6e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OHMSYGLRPGSGSDMSYGLRPGSSQHMYSYGLRPGSSQDMSYGLRPG 49
|||||
DB 927 OHMSYGLRPGSGSDMSYGLRPGSSQHMYSYGLRPGSSQDMSYGLRPG 975
|||||

RESULT 12
US-08-694-865-8
Sequence 8, Application US/08694865
Patent No. 5837268
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO

STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,865
FILING DATE: 09-AUG-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 977 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-865-8

Query Match 100.0%; Score 290; DB 2; Length 977;
Best Local Similarity 100.0%; Pred. No. 1.6e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QHWSYGLRPGSGSDWSYGLRPGSSQHWSYGLRPGSGSDWSYGLRPG 49
DB 927 QHWSYGLRPGSGSDWSYGLRPGSSQHWSYGLRPGSGSDWSYGLRPG 975

RESULT 13
US-08-878-748-8
Sequence 8, Application US/08878748
Patent No. 5969126
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: REDMOND, MARK J.
APPLICANT: HUGHES, HUI P. A.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS
STREET: 635 BRYANT STREET
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,748
FILING DATE: 19-JUN-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: ROBINS, ROBERTA L.

REGISTRATION NUMBER: 33,208
REFERENCE/DOCKET NUMBER: 9001-0016.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 617-8999
TELEFAX: (415) 327-3231
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 977 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-878-748-8

Query Match 100.0%; Score 290; DB 2; Length 977;
Best Local Similarity 100.0%; Pred. No. 1.6e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QHWSYGLRPGSGSDWSYGLRPGSSQHWSYGLRPGSGSDWSYGLRPG 49
DB 927 QHWSYGLRPGSGSDWSYGLRPGSSQHWSYGLRPGSGSDWSYGLRPG 975

RESULT 14
US-09-124-491-8
Sequence 8, Application US/09124491
Patent No. 6022960
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANN, JOHN G.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/124,491
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE: 09-AUG-1996
APPLICATION NUMBER: US 08/387,156
FILING DATE: 10-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/960,932
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,171
FILING DATE: 16-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 977 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-124-491-8

Query Match	100.0%;	Score 290;	DB 3;	Length 977;
Best Local Similarity	100.0%;	Pred. No. 1.6e-25;		
Matches 49;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 QHWSTGLRPGSGSQDWSTYGLRPGSSQHWSTYGLRPGSSQDWSTYGLRPG 49
 |||||
 Db 927 QHWSTGLRPGSGSQDWSTYGLRPGSSQHWSTYGLRPGSSQDWSTYGLRPG 975

RESULT 15
US-07-690-983D-45

GENERAL INFORMATION:

APPLICANT: STEWART, Andrew G

APPLICANT: TSONIS, Con G.

NUMBER OF SEQUENCES:	47
TITLE OF INVENTION:	FUSION PROTEINS
;	
:	

CORRESPONDENCE ADDRESS: NUMBER OF SEQUENCES: 4 /

ADDRESSEE: FOLEY & LA

STREET: 3000 K Street, N.W.

CITY: Washington, D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE
MEDITIM TYPE, E]

MEDIUM TYPE: floppy

OPERATING SYSTEM: PC-DOS

SOFTWARE: PatentIn Release #1.0

```

; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/07/690,983D

FILING DATE: 25-JUN-1991

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: BGT/21190/00373

APPLICATION NUMBER: PCI/A090/003/3
FILING DATE: 24-AUG-1990

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 16786/148

TELECOMMUNICATION INFORMATION:
 TEL: 001-603-633-5300

TELEPHONE: (202)672-5300
TELEFAX: (202)672-5308

1 ELEFAA: (202)672-3399

```

: INFORMATION FOR SEQ ID NO: 43:
:
: SEQUENCE CHARACTERISTICS:
:

```

LENGTH: 44 amino acids

TYPE: amino acid

; TOPOLOGY: linear

```

; MOLECULE TYPE: protein

```

US-07-690-983D-45

Query Match 64.78: Score 18

Best Local Similarity 69.48; Pred. No.

Matches 34; Conservative 4; Misma

64.7%: Score 187.5; DB 1; Length 44;
 Best Local Similarity 69.4%: Pred. No. 2.2e-15;
 Matches 34; Conservative 4; Mismatches 2; Indels 9; Gaps 3

Oy 1 QHWSGLRPGSGSDPWSYGLRPGSSQHWSTYGLRPGSGSDPWSYGLRPG 49
 :|||||: :|||||: :|||||: :|||||:
 db 3 EHWSTGLRPG---EHWSTYGLRPG---EHWSTYGLRPG 42
 :|||||: :|||||: :|||||: :|||||:

Search completed: March 2, 2001, 10:54:37
Job time: 602 sec

THIS PAGE BLANK (USPTO)

OY 46 LRP 48
DB 87 LRP 89

RESULT 2

O9VTR8 PRELIMINARY; PRT: 836 AA.
AC O9VTR8;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ZINC FINGER PROTEIN RIN ZF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99175194; PubMed=10075714;
RA Tiliotson L.G.;
RT "RIN ZF, a novel zinc finger gene, encodes proteins that bind to the
CAC element of the gastrin promoter";
RL J. Biol. Chem. 274:8123-8128(1999).
DR EMBL; AF091457; AAD2522.1; -.
DR HSSP; P07248; IARD.
DR INTERPRO: IPR000210; -.
DR INTERPRO: IPR000822; -.
DR PFAM; PF00096; zf-C2H2; 2.
DR PFAM; PF00651; BTF; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; UNKNOWN_1.
SQ SEQUENCE 836 AA; 91698 MW; E4F48AB2720676C1 CRC64;

Query Match

Best Local Similarity 32.6%; Score 94.5; DB 11; Length 836;
Pred. No. 0.0096;
Matches 22; Conservative 7; Mismatches 14; Indels 3; Gaps 3;

OY 3 WSGLRPGSGSDMSYGLRPGSGSDMSYGLRPG 48
DB 626 FKYGMPGT-SSDFKYGGLP-STSNFKYGGLPGA-PNDFKYGGLP 668

RESULT 3

O9V617 PRELIMINARY; PRT: 1729 AA.
AC O9V617;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE PERICARDINE.
GN PRC.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON R;
RA Chatterjee A., Astler M., Zaffran S., Smeriva M., Gratecos D.;
RT "Characterization of an extracellular matrix component specifically
expressed in the basal lamina of the dorsal vessel in Drosophila
melanogaster";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF203342; AAF13280.1; -.
SQ SEQUENCE 1729 AA; 164661 MW; DA9B1B7FABSCDEDB CRC64;

Query Match

Best Local Similarity 27.9%; Score 81; DB 5; Length 1729;
Pred. No. 0.59; Mismatches 14; Indels 14; Gaps 4;
Matches 22; Conservative 8;

OY 5 YGLRPGSGSDMS-----YGLRPG-----GSSQHMWSYGLRPGSGSDMS-----YGLRPG 49
DB 400 YGSGPGIGGOTGAGCPGYGTNPETIGGOTGAGCP-GVGTDPGICAGTACGAPGYGTGPG 456

RESULT 4

O9VTR6 PRELIMINARY; PRT: 1039 AA.
AC O9VTR6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE CG5700 PROTEIN.
GN CG5700.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Burton R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Plankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Glodek A., Gabrielian A.E., Garg N.S., Galbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Houtlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,
RA Jastali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL; AE003543; AAF49980.1; -.
DR FLYBASE; FBgn0036231; CG5700.
SQ SEQUENCE 1039 AA; 102635 MW; 6A710871139BD01F CRC64;

Query Match

Best Local Similarity 27.8%; Score 80.5; DB 5; Length 1039;
Pred. No. 0.4; Mismatches 15; Indels 11; Gaps 4;
Matches 22; Conservative 6;

OY 5 YGLRPGSGSDMSYGLRPG-----GSSQHMWSYGLRPGSGSDMS-----YGLRPG 49
DB 11 YGLRPGSGSDMSYGLRPG-----GSSQHMWSYGLRPGSGSDMS-----YGLRPG 49

Db 294 YGAPGFGGAP-GYGNOPGVGGGTAGAP-GYGSQPGIGGTAGAGPGYGAOPG 345

```
RESULT 5
ID Q9SFB8 PRELIMINARY; PRT; 715 AA.
AC Q9SFB8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE PUTATIVE PHOSPHATIDYLINOSITOL-4-PHOSPHATE 5-KINASE.
GN F17A17.30.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Cressy T.H., Haas B.,
RA Roming C.M., Koo H., Fujii C.Y., Utechtack T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III BAC F17A17 genomic sequence.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC013483; AAF21206.1; -
DR INTERPRO: IPR000169; -
DR INTERPRO: IPR002498; -
DR PFAM: PF01504; PIP5K; 1.
DR PROSITE: PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
KW Kinase.
SQ SEQUENCE 715 AA; 81422 MW; 795D86CEB1FB52E9 CRC64;
```

Query Match 27.4%; Score 79.5; DB 10; Length 715;
Best Local Similarity 41.9%; Pred. No. 0.35;
Matches 18; Conservative 4; Mismatches 18; Indels 3; Gaps 1;

OY 3 WSYGLRPGSSQDMSYGLRPGSSQHSYGLRPGSSQDWSYG 45
DB 151 WKNGLMNGMTIMSNGRYDGS--WEDGAPKRGKGTFRWSDG 190

```
RESULT 6
ID 000509 PRELIMINARY; PRT; 197 AA.
AC 000509;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE HYPOHETICAL 21.0 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA, AND PBM;
RX MEDLINE=96125193; PubMed=8543165;
RA Gualandi G., Frezza D., Scotto A., Bianchi E., Gargano S.,
RA Fruscalzo A., Calef E.;
RT "Integration of an Epstein-Barr virus episome 3' into the gene
RT encoding immunoglobulin heavy-chain alpha 1 in a lymphoblastoid cell
RT line.";
RL Gene 166:221-226(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA, AND PBM;
RA Gualandi G.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA, AND PBM;
```

```
RA Gualandi G.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA, AND PBM;
RA Frezza D.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA, AND PBM;
RA Frezza D.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: X76785; CA54181.1; -
KW Hypothetical protein.
SQ SEQUENCE 197 AA; 20977 MW; 7EF9725ED9D18CCF CRC64;
```

Query Match 26.9%; Score 78; DB 4; Length 197;
Best Local Similarity 27.7%; Pred. No. 0.14;
Matches 23; Conservative 5; Mismatches 21; Indels 34; Gaps 3;

OY 1 QHWSYGLRPGSG-----SQDMSYGLRPGSG-----SQHWSYGLRPGSG----- 38
DB 61 QWSSSLSPGRHSALDLVTQPSWSSLSPPRHSALVLTQPSWSSLSPPRHSALDLIT 120
OY 39 -----SQDMSYGLRPG 49
DB 121 HPGPHALDLITQPSWSSLSPG 143

```
RESULT 7
ID 09M815 PRELIMINARY; PRT; 769 AA.
AC 09M815;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE PUTATIVE PHOSPHATIDYLINOSITOL-4-PHOSPHATE 5-KINASE.
GN T7P1.4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Cressy T.H., Haas B., Wu D.,
RA Roming C.M., Koo H., Fujii C.Y., Utechtack T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome I BAC T7P1 genomic sequence.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC018908; AAF27664.1; -
KW Kinase.
SQ SEQUENCE 769 AA; 87585 MW; 652F441B2F493845 CRC64;
```

Query Match 26.2%; Score 76; DB 10; Length 769;
Best Local Similarity 29.9%; Pred. No. 0.91;
Matches 23; Conservative 4; Mismatches 20; Indels 30; Gaps 2;

OY 3 WSYGLRPGSSQDMSYGLRPGSG-----SQHWSYGLRPGSGSG-- 40
DB 116 WREGLDGSSGSWYNGRFTGNMKKMGSRGVSWANGDLFNGFWLNGLRHSGGYKY 175

OY 41 -----DMSYGLRPG 49
DB 176 ADGCFYFGTWSRGLKDG 192

```
RESULT 8
O97643 PRELIMINARY; PRT; 462 AA.
ID O97643
```

AC 097643;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
 DE FIBRINOGEN A ALPHA CHAIN (FRAGMENT).
 OS Lama glama (Llama).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
 OX NCBI_TaxID=9844;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Murakawa M.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF095464; AAC67563.1; -
 DR HSSP; P02671; 1FZG.
 FT NON_TER 1 1
 FT NON_TER 462 462
 SQ SEQUENCE 462 AA; 48968 MW; 377B503CF32E698 CRC64;

Query Match
 Best Local Similarity 25.3%; Score 73.5; DB 6; Length 462;
 Matches 19; Conservative 9; Mismatches 17; Indels 5; Gaps 5;

QY 3 WSYGL-RPGSGSDMSYGLRPGSSQHSYGL-RPSSGSDMSYGL-RPG 49
 Db 166 WTHGSEPPGSAG-TWTHGSEPPGSAGWTHGLEPPGSAG-SWTHGSEPPG 213

RESULT 9
 ID 013344 PRELIMINARY; PRT; 528 AA.
 AC 013344;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE FUS-LIKE PROTEIN (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tissue-Carcinoma;
 RA Itch K., Kawase M.;
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROBABLE RNA-BINDING PROTEIN (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RNP).
 DR EMBL; U36561; AAA79948.1; -
 DR INTERPRO: IPR000504; -
 DR INTERPRO: IPR001419; -
 DR INTERPRO: IPR001876; -
 DR INTERPRO: IPR002961; -
 DR PFM; PFO0076; rtm; 1.
 DR PFM; PFO0641; zf-RanBP; 1.
 DR PRINTS: PRO00210; GLUTENIN.
 DR PRINTS: PRO1237; TNFC.
 DR PROSITE: PS00030; RNP_1; FALSE_NEG.
 KW RNA-binding; Nuclear protein.
 FT DOMAIN 1 180
 FT DOMAIN 159 164
 FT DOMAIN 163 168
 FT DOMAIN 179 198
 FT DOMAIN 181 270
 FT DOMAIN 205 234
 FT DOMAIN 226 234
 FT DOMAIN 289 294
 FT DOMAIN 337 344
 FT DOMAIN 374 527
 FT DOMAIN 396 412
 FT DOMAIN 422
 POLY-SER.
 POLY-GLY.
 POLY-GLY.
 GLY-RICH.
 POLY-GLY.
 POLY-GLY.
 RNA-BINDING (RNP2) (BY SIMILARITY).
 RNA-BINDING (RNP1) (BY SIMILARITY).
 ARG/GLY-RICH.
 POLY-GLY.
 POLY-GLY.

FT ZN_FING 431 450 C4-TYPE (POTENTIAL).
 FT NON_TER 528
 SQ SEQUENCE 528 AA; 53377 MW; 43CFD6CBDA57A258 CRC64;

Query Match
 Best Local Similarity 24.8%; Score 72; DB 4; Length 528;
 Matches 20; Conservative 2; Mismatches 12; Indels 16; Gaps 2;

QY 5 YGLRPGSDMSYGLRPGSSQHSYGLRPGSDMSYGLRPG 45
 Db 99 YGPQPTSTSGSYSSQTSSTGQPP-----STGQPPSTGQDQDSTG 141

RESULT 10
 ID 091020 PRELIMINARY; PRT; 701 AA.
 AC 091020;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE HYPOTHETICAL 73.3 KDA PROTEIN.
 GN SCL6.26C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Redenbach M., Kieser H.M., Denaplatte D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL MOL. Microbiol. 21:77-96(1996).
 DR EMBL; AL159139; CAB76891.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 701 AA; 73279 MW; 558DFCBCT77BC412 CRC64;

Query Match
 Best Local Similarity 24.8%; Score 72; DB 2; Length 701;
 Matches 21; Conservative 4; Mismatches 19; Indels 10; Gaps 3;

QY 5 YGLRPGSDMSYGLRPG-----GSSQHS---YGLRPGSDMSYGLRPG 49
 Db 577 HGARDHGGRDGDGARGPHGTGDDGGDHRSDGARGDYGARD-GYGAODG 629

RESULT 11
 ID 057480 PRELIMINARY; PRT; 1068 AA.
 AC 057480;
 DT 01-JUN-1998 (TReMBLrel. 06, Created)
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE PHOTORECEPTOR GUANYLATE CYCLASE 1 (FRAGMENT).
 GN GCL.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RHODE ISLAND RED; TISSUE-RETINA-PIGMENT EPITHELIUM-CHOROIO;
RA Semple-Rowland S.L., Lee N.R., Van Hooser J.P., Palczewski K.,
  Baehr W.;
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(1997).
DR EMBL: AF036942; AAC24500.1; -.
DR HSSP: 002846; IAML.
DR INTERPRO: IPR000719; -.
DR INTERPRO: IPR001054; -.
DR INTERPRO: IPR001245; -.
DR INTERPRO: IPR001828; -.
DR PFAM: PF00069; Kinase; 1.
DR PFAM: PF00211; guanylate cyc; 1.
DR PFAM: PF01094; ANF_receptor; 1.
DR PRINTS: PR00109; TRYKINASE.
DR PROSITE: PS00452; GUANYLATE_CYCLASES; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KM Lyase.
FT NON_TER
SQ SEQUENCE 1068 AA; 11828 MW; 3FAC693E1597174B CRC64;

Query Match 24.8%; Score 72; DB 13; Length 1068;
Best Local Similarity 39.0%; Pred. No. 3.4;
Matches 16; Conservative 3; Mismatches 14; Indels 8; Gaps 2;

OY 10 GSGSDMS-YGLRPG-----GSSQHSYGLRPGSGSDM 42
DB 356 GKGDRLMPVGLPEIRGLAYRGHSHVHPHSSPGTDSGCW 396

RESULT 12
OY 09T513 PRELIMINARY; PRT; 107 AA.
AC 09T513;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE CHICKEN LUTEINIZING HORMONE-RELEASING HORMONE II (FRAGMENT).
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Abler L.A., Sherwood N.M., Grendell R.L., Golos T.G., Terasawa E.;
RT "CDNA of a second form of luteinizing hormone releasing, chicken LHRH-
  II, isolated from the non-human primate brain (Abstract 632.8).";
RL Abstr. Soc. Neurosci. 24:1607-1607(1998).
DR EMBL: AF104307; AADI3775.1; -.
DR INTERPRO: IPR002012; -.
DR PFAM: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
FT NON_TER
SQ SEQUENCE 107 AA; 11823 MW; FACE52703C3CB1D CRC64;

Query Match 24.5%; Score 71; DB 6; Length 107;
Best Local Similarity 46.2%; Pred. No. 0.44;
Matches 18; Conservative 3; Mismatches 12; Indels 6; Gaps 2;

OY 1 QHWSYGLRPG-----SGSDMSYGLR-PGSSQHSYGL 33
DB 25 QHWSHGWPYGGKRALSSADPQNALRPPAGSPAATYGL 63

RESULT 13
OY 097655 PRELIMINARY; PRT; 114 AA.
AC 097655;
DT 01-MAY-1999 (Tremblrel. 10, Created)

```

```

DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DE 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LULIBERIN)
DE PRECURSOR.
GN GNRH2.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA White R.B., Urbanski H.F., Fernald R.D.;
RT "A second gene for gonadotropin-releasing hormone is expressed in the
  rhesus macaque (Abstract #632.18).";
RL Abstr. Soc. Neurosci. 24:1609-1609(1998).
CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
  POLICIE-STIMULATING HORMONES.
CC -1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
DR EMBL: AF097356; AAD09106.1; -.
DR INTERPRO: IPR002012; -.
DR PFAM: PF00446; GnRH; 1.
DR PROSITE: PS00473; GnRH; 1.
DR PRODOM: PD029787; -. 1.
KM Hormone; Amidation.
FT NON_TER
SQ SEQUENCE 114 AA; 12533 MW; 8B70D690D5BD5103 CRC64;

Query Match 24.5%; Score 71; DB 6; Length 114;
Best Local Similarity 46.2%; Pred. No. 0.47;
Matches 18; Conservative 3; Mismatches 12; Indels 6; Gaps 2;

OY 1 QHWSYGLRPG-----SGSDMSYGLR-PGSSQHSYGL 33
DB 25 QHWSHGWPYGGKRALSSADPQNALRPPAGSPAATYGL 63

RESULT 14
OY 09X111 PRELIMINARY; PRT; 257 AA.
AC 09X111;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE F19G14.22 PROTEIN.
GN F19G14.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Strain-CV, COLUMBIA;
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
  Barnstead M.E., Mason T.M., Bowman C.L., Rongning C.M., Benito M.,
  Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
  Fraser C.M., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC F19G14 genomic sequence.";
RL submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC006438; AAD11986.1; -.
DR HSSP: P00303; 2CBP.
FT NON_TER
SQ SEQUENCE 257 AA; 27400 MW; DA59B77A93D53A17 CRC64;

Query Match 24.5%; Score 71; DB 10; Length 257;
Best Local Similarity 26.1%; Pred. No. 1.1;
Matches 18; Conservative 6; Mismatches 19; Indels 26; Gaps 3;

OY 3 WSYGLRPG-----SGSDMSYGLR-PGSSQHSYGLRPGS----- 37
DB 87 WGMGVVPMNTHSSGSGGCMGCPNNNYSGSGSGSGGSMGVGSHSKNVMNATYNGPKRIIV 146

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using 'sw model

Run on: March 2, 2001, 11:01:20 ; Search time 32.03 Seconds
(Without alignments)
49,404 Million cell updates/sec

Title: US-09-306-689-11

Perfect score: 290
Sequence: 1 QHWSYGLRPGSGSDMSYGL.....SYGLRPGSGSDMSYGLRPG 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SWISSPROT_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	77	27.2	407	1 SM41_HEMPU	Q26264 hemilectrot
2	79	26.6	445	1 SM50_SRP	P16994 strongyloce
3	76.5	26.4	67	1 GONI_MACMU	P55247 macaca mula
4	76.5	26.4	92	1 GONI_HUMAN	P01148 homo sapien
5	74	25.5	90	1 GONI_MOUSE	P13562 mus musculu
6	74	25.5	91	1 GONI_PIG	P49921 sus scrofa
7	74	25.5	92	1 GONI_RAT	P07490 rattus norv
8	71	24.5	414	1 SVS2_RAT	P22006 rattus norv
9	70	24.1	518	1 FUS_MOUSE	P56959 mus musculu
10	69	23.8	512	1 FUS_BOVIN	Q28009 bos taurus
11	69	23.8	526	1 FUS_HUMAN	P35637 homo sapien
12	68.5	23.6	94	1 GONI_HABBU	P51918 haplochromi
13	68	23.4	335	1 SM34_LYTP1	Q05994 lytechinus
14	66.5	22.9	89	1 GONI_XENLA	P45656 xenopus lae
15	66	22.8	776	1 ISOA_PSEAY	P10342 pseudomonas
16	66	22.8	776	1 ISOA_PSEAY	P26501 pseudomonas
17	65	22.4	92	1 GONI_TURGB	Q95335 cupata glis
18	65	22.4	120	1 GONI_HUMAN	Q43555 homo sapien
19	64	22.1	92	1 GONI_CAVPO	O54752 cavia porce
20	63.5	21.9	95	1 GONI_PAGMA	P70074 pagrus majo
21	63.5	21.9	95	1 GONI_SPAUA	P51919 sparus auro
22	63	21.7	61	1 GONI_SHEEP	Q28588 ovis aries
23	63	21.7	63	1 GONI_MESAU	Q09183 mesocricetu
24	62	21.4	839	1 GLUT5_WHEAT	P10368 tritlicum ae
25	61	21.0	767	1 ACES_CHICK	P36196 gallus galli
26	60.5	20.9	627	1 K2C1_MOUSE	P41044 mus musculu
27	60	20.7	92	1 GONI_CHICK	P77042 gallus galli
28	60	20.7	379	1 ROA3_HUMAN	P51991 homo sapien
29	60	20.7	505	1 SCRY_SALTY	P22340 salmoneilla
30	59.5	20.3	838	1 GLUT4_WHEAT	P08489 tritlicum ae
31	59	20.3	10	1 GONI_ALMTI	P77041 alligator m
32	59	20.3	114	1 GONI_TURGB	Q95336 cupata glis
33	59	20.3	567	1 ENV_AVISN	P31796 avian splee

34	58	20.0	80	1 GONI_CLAGA	P33439 ciarias gar
35	58	20.0	82	1 GONI_SALTR	P45653 salmo truit
36	58	20.0	370	1 ALIS_DIPPU	P12764 diptoptera
37	57	19.7	385	1 HEM6_SOTBN	P35055 glycine max
38	57	19.7	391	1 HEM6_HORVU	Q42840 hordeum vul
39	57	19.7	492	1 SYNW_YEAST	P25345 saccharomyc
40	57	19.7	1321	1 IRS2_MOUSE	P81122 mus musculu
41	56	19.3	219	1 TRPF_CAUCR	P12289 caulobacter
42	56	19.3	313	1 SYPH_HUMAN	P08247 homo sapien
43	56	19.3	373	1 RO31_XENLA	P51968 xenopus lae
44	56	19.3	385	1 RO32_XENLA	P51992 xenopus lae
45	55.5	19.1	86	1 GONI2_CLAGA	P43306 ciarias gar

ALIGNMENTS

```

RESULT 1
ID SM41_HEMPU STANDARD; PRT; 407 AA.
AC Q26264;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 41 KDA SPICULE MATRIX PROTEIN PRECURSOR (HSM41) (HPSMC).
OS Hemilectrotus pulcherrimus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinozoa; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinozoa; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinozoa; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Hemilectrotus.
RN [1]
RP MEDLINE=93075615; PubMed=1445780;
RA Katoh-Fukui Y., Noce T., Ueda T., Fujiwara Y., Hashimoto N.,
RA Tanaka S., Higashinakagawa T.;
RT *Isolation and characterization of cDNA encoding a spicule matrix
RT protein in Hemilectrotus pulcherrimus micromeres. ";
RL Int. J. Dev. Biol. 36:353-361(1992).
CC -1- FUNCTION: MAJOR MATRIX PROTEIN OF THE SEA URCHIN EMBRYO SPICULE
CC WHICH DIRECTS CRYSTAL GROWTH IN CERTAIN ORIENTATIONS AND INHIBIT
CC GROWTH IN OTHERS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE
CC MICROMERE/PRIMARY MESENCHYME CELLS (PMC) LINEAGE.
CC -1- DEVELOPMENTAL STAGE: APPEARS IN GASTRULAE AND REMAINS AT A SIMILAR
CC LEVEL UNTIL THE PLUTEUS STAGE.
CC -1- DOMAIN: THE REPETITIVE DOMAIN MAY PROVIDE A CALCITE BINDING
CC MATRIX.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE SM50 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collapration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: S48755; AAB24285.1; -.
CC INTERPRO: IPR001304; -.
CC DR PROSITE: PS00615; C-TYPE LECTIN_1; FALSE_NEG.
CC PROSITE: PS50041; C-TYPE LECTIN_2; 1.
CC DR Matrix protein; Signal; Repeat.
CC FT SIGNAL 1 17 POTENTIAL.
CC FT CHAIN 18 407 41 KDA SPICULE MATRIX PROTEIN.
CC FT DOMAIN ? ? C-TYPE LECTIN.
CC FT DOMAIN 104 179 PRO-RICH.
CC SO SEQUENCE 407 AA; 42383 MM; 5A5397C46B80F6B5 CRC64;

Query Match 27.2%; Score 79; DB 1; Length 407;
Best Local Similarity 30.6%; Pred. No. 0.15;
Matches 15; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

```


Db 2 GCSSQHSYGLRPG-GRKD 19

RESULT 4

ID	GONI_HUMAN	STANDARD	PRT	92 AA
AC	P01148			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-APR-1988 (Rel. 07, Last sequence update)			
DE	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I) (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING HORMONE I) (GNRH I) (LULIBERIN I) (GONADORELIN) (GNRH-ASSOCIATED PEPTIDE I)]			
GN	GNRH1 OR GNRH OR LHRH			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominoidea; Homo			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE-89366682; PubMed-2671939;			
RT	Hayflick J.S., Adelman J.P., Seeburg P.H.;			
RT	"The complete nucleotide sequence of the human gonadotropin-releasing hormone gene."			
RL	Nucleic Acids Res. 17:6403-6403(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE-86094338; PubMed-2867548;			
RT	Adelman J.P., Mason A.J., Hayflick J.S., Seeburg P.H.;			
RT	"Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin-releasing hormone and prolactin release-inhibiting factor in human and rat."			
RL	Proc. Natl. Acad. Sci. U.S.A. 83:179-183(1986).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE-85012739; PubMed-6090951;			
RT	Seeburg P.H., Adelman J.P.;			
RT	"Characterization of cDNA for precursor of human luteinizing hormone releasing hormone."			
RL	Nature 311:666-668(1984).			
RN	[4]			
RP	SEQUENCE OF 24-33.			
RA	MEDLINE-83126573; PubMed-6760865;			
RT	Tan L., Rousseau P.;			
RT	"The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in the human placenta."			
RL	Biochem. Biophys. Res. Commun. 109:1061-1071(1982).			
CC	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.			
CC	-1- PHARMACEUTICAL: AVAILABLE UNDER THE NAMES FACTREL (AYERST LABS), LUTREPUSE OR LUTRELER (FERRING PHARMACEUTICALS) AND RELISORM (SERONO).			
CC	-1- SIMILARITY: BELONGS TO THE GNRH FAMILY.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: X01059; CAA25526.1; -			
DR	EMBL: M12578; AAA35916.1; -			
DR	EMBL: X15215; CAA33285.1; -			
DR	PIR: A01410; RHUG.			
DR	PIR: A26173; A26173.			
DR	PIR: S05308; S05308.			
DR	MTM: 152760; -			
DR	INTERPRO: IPR002012; -			
DR	PFAM: PF00446; GNRH.1.			
DR	PROSITE: PS00473; GNRH.1.			

KW Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Placenta; Pharmaceutical; Signal.

FT	SIGNAL	1	23	
FT <td>CHAIN</td> <td>24</td> <td>92</td> <td>PROGONADOLIBERIN I.</td>	CHAIN	24	92	PROGONADOLIBERIN I.
FT <td>PEPTIDE</td> <td>24</td> <td>33</td> <td>GONADOLIBERIN I.</td>	PEPTIDE	24	33	GONADOLIBERIN I.
FT <td>PEPTIDE</td> <td>37</td> <td>92</td> <td>GNRH-ASSOCIATED PEPTIDE I.</td>	PEPTIDE	37	92	GNRH-ASSOCIATED PEPTIDE I.
FT <td>ACT_SITE</td> <td>26</td> <td>26</td> <td>APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.</td>	ACT_SITE	26	26	APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.
FT <td>MOD_RES</td> <td>24</td> <td>24</td> <td>PYRROLIDONE CARBOXYLIC ACID.</td>	MOD_RES	24	24	PYRROLIDONE CARBOXYLIC ACID.
FT <td>MOD_RES</td> <td>33</td> <td>33</td> <td>AMIDATION (G-34 PROVIDE AMIDE GROUP).</td>	MOD_RES	33	33	AMIDATION (G-34 PROVIDE AMIDE GROUP).
FT <td>CONFLICT</td> <td>16</td> <td>16</td> <td>W -> S (IN REF. 3).</td>	CONFLICT	16	16	W -> S (IN REF. 3).
SO <td>SEQUENCE</td> <td>92 AA;</td> <td>10380 MW;</td> <td>30A72221B076FA79 CRC64;</td>	SEQUENCE	92 AA;	10380 MW;	30A72221B076FA79 CRC64;

Query Match Score 76.5; DB 1; Length 92;
Best Local Similarity 78.96; Pred. No. 0.063;
Matches 15; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Db 20 GCSSQHSYGLRPG-GRKD 37

RESULT 5

ID	GONI_MOUSE	STANDARD	PRT	90 AA
AC	P13562			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	01-JAN-1990 (Rel. 13, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	PROGONADOLIBERIN I PRECURSOR [CONTAINS: GONADOLIBERIN I (LHRH I) (LUTEINIZING HORMONE RELEASING HORMONE I) (GONADOTROPIN RELEASING HORMONE I) (GNRH I) (LULIBERIN I); PROLACTIN RELEASE-INHIBITING FACTOR I]			
GN	GNRH1 OR GNRH			
OS	Mus musculus (Mouse)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Scurionath; Muridae; Murinae; Mus			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE-87069928; PubMed-3024317;			
RA	Mason A.J., Hayflick J.S., Zoeller R.T., Young W.S. III,			
RA	Phillips H.S., Nikolics K., Seeburg P.H.;			
RT	"A deletion truncating the gonadotropin-releasing hormone gene is responsible for hypogonadism in the hpg mouse."			
RL	Science 234:1366-1371(1986).			
CC	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING HORMONES.			
CC	-1- SIMILARITY: BELONGS TO THE GNRH FAMILY.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: M14872; AAA37717.1; -			
DR	MGD: MGI:95789; GNRH.			
DR	INTERPRO: IPR002012; -			
DR	PFAM: PF00446; GNRH.1.			
DR	PROSITE: PS00473; GNRH.1.			
KW	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus; Placenta; Signal.			
FT <td>SIGNAL</td> <td>1</td> <td>21</td> <td></td>	SIGNAL	1	21	
FT <td>CHAIN</td> <td>22</td> <td>90</td> <td>PROGONADOLIBERIN I.</td>	CHAIN	22	90	PROGONADOLIBERIN I.
FT <td>PEPTIDE</td> <td>22</td> <td>31</td> <td>GONADOLIBERIN I.</td>	PEPTIDE	22	31	GONADOLIBERIN I.
FT <td>PEPTIDE</td> <td>35</td> <td>90</td> <td>PROLACTIN RELEASE-INHIBITING FACTOR I.</td>	PEPTIDE	35	90	PROLACTIN RELEASE-INHIBITING FACTOR I.
FT <td>ACT_SITE</td> <td>24</td> <td>24</td> <td>APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.</td>	ACT_SITE	24	24	APPEARS TO BE ESSENTIAL FOR BIOLOGICAL ACTIVITY.
FT <td>MOD_RES</td> <td>22</td> <td>22</td> <td>PYRROLIDONE CARBOXYLIC ACID.</td>	MOD_RES	22	22	PYRROLIDONE CARBOXYLIC ACID.

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: S50870; AAB24572.1; -
DR EMBL: M12579; AAA41263.1; -
DR EMBL: M31670; AAA41264.1; -
DR EMBL: M15527; AAA42141.1; ALT_SEQ.
DR EMBL: M15529; AAA42139.1; -
DR EMBL: M15528; -; NOT_ANNOTATED_CDS.
DR PIR: B26173; RHRTG.
DR PIR: A48410; A48410.
DR INTERPRO: IPR002012; -
DR PRAM: PF00446; GNRH; 1.
DR PROSITE: PS00473; GNRH; 1.
KM Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KW Placenta; Signal.

FT SIGNAL 1 23
FT CHAIN 24 92 PROCONADOLIBERIN I.
FT PEPTIDE 24 33 GONADOLIBERIN I.
FT PEPTIDE 37 92 PROLACTIN RELEASE-INHIBITING FACTOR I.
FT ACT_SITE 26 26 APPEARS TO BE ESSENTIAL FOR BIOLOGICAL
ACTIVITY.
FT MOD_RES 24 24 PYROLIDONE CARBOXYLIC ACID.
FT MOD_RES 33 33 AMIDATION (G-34 PROVIDE AMIDE GROUP).
SQ SEQUENCE 92 AA; 10500 MW; 494B5C64DA8A3EB3 CRC64;

QY 23 GGSQSHWSYGLRPG 36
ID 1 | | | | | | | | | |
Db 20 GCSSQSHWSYGLRPG 33

Query Match 25.5%; Score 74; DB 1; Length 92;
Best Local Similarity 92.9%; Pred. No. 0.11;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 8
SVS2_RAT STANDARD: PRT; 414 AA.
AC P22006;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE SEMINAL VESICLE SECRETORY PROTEIN II PRECURSOR (SVS II).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90277684; Pubmed=2351680;
RA Harris S.E., Harris M.A., Johnson C.M., Bean M.F., Dodd J.G.,
RA Mathus R.J., Cair S.A., Crabh J.W.;
RT "Structural characterization of the rat seminal vesicle secretion II
protein and gene."
RL J. Biol. Chem. 265:9896-9903(1990).
CC -1- FUNCTION: THE RAT SEMINAL VESICLE CONTAINS SIX MAJOR ANDROGEN-
DEPENDENT SECRETORY PROTEINS REFERRED TO AS SVS I-VI. THE SVS
CC I-III PROTEINS APPEAR TO BE COMPONENTS OF THE RAT COPULATORY
CC PLUG. WITH THE SVS II PROTEIN BEING THE MAJOR COMPONENT.
CC -1- PTM: THE REPEATING UNIT APPEARS TO BE INVOLVED IN THE FORMATION OF
CC THE COPULATORY PLUG VIA A TRANSGLUTAMINASE REACTION CROSS-LINKING
CC GLUTAMINE AND LYSINE RESIDUES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC or send an email to license@isb-sib.ch).

DR EMBL: J05443; AAA42192.1; -
DR PIR: A36443; A36443.
DR INTERPRO: IPR002080; -
DR PROSITE: PS00515; SVP_II; 12.
KW Testosterone; Seminal vesicle; Signal; Repeat; Copulatory plug.

FT SIGNAL 1 22
FT CHAIN 23 414 SEMINAL VESICLE SECRETORY PROTEIN II.
FT MOD_RES 23 23 PYROLIDONE CARBOXYLIC ACID.
FT DOMAIN 108 311 13 x 13 AA TANDEM REPEATS.
FT REPEAT 108 120 1.
FT REPEAT 127 139 2.
FT REPEAT 140 152 3.
FT REPEAT 153 165 4.
FT REPEAT 166 178 5.
FT REPEAT 179 191 6.
FT REPEAT 192 204 7.
FT REPEAT 205 217 8.
FT REPEAT 224 236 9.
FT REPEAT 237 249 10.
FT REPEAT 257 269 11.
FT REPEAT 275 287 12.
FT REPEAT 299 311 13.
SQ SEQUENCE 414 AA; 45539 MW; C89E3FCE0CDE92E CRC64;

QY 4 SYGLRPGSGSDMSYGLRPGSSQSHWSYGLRPGSGSDMSG 45
ID 1 | | | | | | | | | |
Db 184 SFGQMKSSGSQVKSFGQMKASESQIKSFQKRSQGLQSYG 225

Query Match 24.5%; Score 71; DB 1; Length 414;
Best Local Similarity 42.9%; Pred. No. 0.99;
Matches 18; Conservative 3; Mismatches 21; Indels 0; Gaps 0;

RESULT 9
FUS_MOUSE STANDARD: PRT; 518 AA.
AC P56959;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE RNA-BINDING PROTEIN FUS (PIGEPN PROTEIN).
GN FUS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA Alappat S.R., Zhang M., Zhao X., Allegro M.A., Allegro M.C.,
RA Burdall C.A.;
RT "Regulation of pigen expression in mouse embryos."
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: BINDS BOTH SINGLE-STRANDED AND DOUBLE-STRANDED DNA AND
CC PROMOTES ATP-INDEPENDENT ANNEALING OF COMPLEMENTARY SINGLE-
CC STRANDED DNAs AND D-LOOP FORMATION IN SUPERHELICAL DOUBLE-STRANDED
CC DNA. MAY PLAY A ROLE IN MAINTENANCE OF GENOMIC INTEGRITY (BY
CC SIMILARITY).
CC -1- SUBUNIT: COMPONENT OF NUCLEAR RIBOPROTEIN COMPLEXES (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF224264; AAF70602.1; -
 DR MGD: MGI:1353633; FUS.
 DR PROSITE: PS50102; RRM; 1.
 DR RNA-binding; DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;
 DR Metal-binding.
 KW DOMAIN 1 167 GLN/GLY/SER/TYR-RICH.
 FT DOMAIN 168 265 GLY-RICH.
 FT DOMAIN 278 364 RNA-BINDING (RRM).
 FT DOMAIN 364 518 ARG/GLY-RICH.
 FT ZN_FING 421 440 CA-TYPE (POTENTIAL).
 FT ZN_FING 518 AA; 52673 MW; 306F231BEED78D6 CRC64;
 SQ SEQUENCE

Query Match 24.1%; Score 70; DB 1; Length 518;
 Best Local Similarity 40.8%; Pred. No. 1.6; Mismatches 15; Indels 12; Gaps 2;
 Matches 20; Conservative 2; Mismatches 15; Indels 12; Gaps 2;

OY 5 YGLRPGSGSDMSYGLRPGSSQHWMS-----YGLRPGSGSDMSYSG 45
 Db 101 YGQGPAPSSSTSGSY-----GSSSQSSSYGDPQSGGYGQSGYGGQOQSYG 145

RESULT 10
 FUS_BOVIN STANDARD; PRT; 512 AA.
 AC 028009;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE RNA-BINDING PROTEIN FUS (PIGEPEN PROTEIN).
 OS Bos taurus (Bovine).
 GN FUS.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-AORTA;
 RX MEDLINE-96175600; PubMed-8631501;
 RA Allegro M.C.; Allegro M.A.;
 RT "A nuclear protein regulated during the transition from active to
 RT quiescent phenotype in cultured endothelial cells."
 RL Dev. Biol. 174:288-297(1996).
 RN [2]
 RP CARBOHYDRATE BINDING DOMAIN
 RX MEDLINE-20160719; PubMed-10694442;
 RA Allegro M.C.;
 RT "A C-terminal carbohydrate-binding domain in the endothelial cell
 RT regulatory protein, pigpen: new function for an EWS family member."
 RL Exp. Cell Res. 255:270-277(2000).
 CC -1- FUNCTION: BINDS BOTH SINGLE-STRANDED AND DOUBLE-STRANDED DNA AND
 CC PROMOTES ATP-INDEPENDENT ANNEALING OF COMPLEMENTARY SINGLE-
 CC STRANDED DNAs AND D-LOOP FORMATION IN SUPERHELICAL DOUBLE-STRANDED
 CC DNA. MAY PLAY A ROLE IN MAINTENANCE OF GENOMIC INTEGRITY (BY
 CC SIMILARITY).
 CC -1- SUBUNIT: COMPONENT OF NUCLEAR RIBOPROTEIN COMPLEXES (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR, EXHIBITS DIFFUSE STAINING
 CC THROUGHOUT (EXCLUDING NUCLEOLI), TOGETHER WITH A SMALL NUMBER OF
 CC INTENSELY STAINED FOCAL POINTS, OR GRANULES, AND PUNCTATE STAINING
 CC ALONG THE NUCLEAR ENVELOPE.
 CC -1- DOMAIN: THE C-TERMINAL DOMAIN BINDS CARBOHYDRATES.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U26024; AAC13543.1; -
 DR HSSP: P09651; 1HA1.
 DR INTERPRO: IPR000504; -
 DR INTERPRO: IPR001876; -
 DR PFAM: PF00076; rtm; 1.
 DR PFAM: PF00641; zt-RanBP; 1.
 DR PROSITE: PS50102; RRM; 1.
 DR PROSITE: PS00030; RRM_RNP_1; FALSE NEG.
 KW RNA-binding; DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;
 KW Metal-binding.
 FT DOMAIN 1 164 GLN/GLY/SER/TYR-RICH.
 FT DOMAIN 165 253 GLY-RICH.
 FT DOMAIN 271 357 RNA-BINDING (RRM).
 FT DOMAIN 357 512 ARG/GLY-RICH.
 FT ZN_FING 414 433 CA-TYPE (POTENTIAL).
 FT ZN_FING 512 AA; 52240 MW; 3652329C044F1386 CRC64;
 SQ SEQUENCE

Query Match 23.8%; Score 69; DB 1; Length 512;
 Best Local Similarity 38.0%; Pred. No. 1.9; Mismatches 15; Indels 14; Gaps 2;
 Matches 19; Conservative 2; Mismatches 15; Indels 14; Gaps 2;

OY 5 YGLRP-----GSGSDMSYGLRPGSSQHWMSYGLRPGSGSDMSYSG 45
 Db 100 YGQGPAPSSSTSGSYGSSSQSSSYGDPQGC-----GYGQSGYGGQOQSYG 144

RESULT 11
 FUS_HUMAN STANDARD; PRT; 526 AA.
 AC P35637;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE RNA-BINDING PROTEIN FUS (ONCOGENE FUS) (ONCOGENE TLS) (TRANSLOCATED IN
 DE LIPOSARCOMA PROTEIN) (POMP75) (75 KDA DNA-PAIRING PROTEIN).
 GN FUS OR TLS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RX MEDLINE-93288139; PubMed-8510758;
 RA Crozat A., Aman P., Mandahl N., Ron D.;
 RT "Fusion of CHOP to a novel RNA-binding protein in human myxoid
 RT liposarcoma."
 RL Nature 363:640-644(1993).
 RN [2]
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).
 RX MEDLINE-93350637; PubMed-7503811;
 RA Rablitt T.H., Forster A., Larson R., Nathan P.;
 RT "Fusion of the dominant negative transcription regulator CHOP with a
 RT novel gene FUS by translocation t(12;16) in malignant liposarcoma."
 RL Nat. Genet. 4:175-180(1993).
 RN [3]
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RX MEDLINE-99013873; PubMed-9795213;
 RA Morohoshi F., Ootsuka Y., Arai K., Ichikawa H., Mitani S.,
 RA Munakata N., Ohki M.;
 RT "Genomic structure of the human RBP56/hRAF168 and FUS/TLS genes."
 RL Gene 221:191-198(1998).
 RN [4]
 RP SEQUENCE OF 235-244; 307-312; 335-345 AND 349-357, & CHARACTERIZATION.
 RX MEDLINE-20036580; PubMed-10567410;
 RA Baechtold H., Kuroda M., Sok J., Kon D., Lopez B.S., Akhmedov A.T.;
 RT "Human 75-kDa DNA-pairing protein is identical to the pro-oncoprotein
 RT TLS/FUS and is able to promote D-loop formation."
 RL J. Biol. Chem. 274:34337-34342(1999).
 RN [5]
 RP CHROMOSOMAL TRANSLOCATION.
 RX MEDLINE-94243799; PubMed-8187069;
 RA Ichikawa H., Shimizu K., Hayashi Y., Ohki M.;
 RT "An RNA-binding protein gene, TLS/FUS, is fused to ERG in human

```

RT      myeloidleukemia with t(16;21) chromosomal translocation." ;
RL      Cancer Res. 54:2865-2868(1994)
CC      -1- FUNCTION: BINDS BOTH SINGLE-STRANDED AND DOUBLE-STRANDED DNA AND
CC      PROMOTES ATP-INDEPENDENT ANNEALING OF COMPLEMENTARY SINGLE-
CC      STRANDED DNAs AND D-LOOP FORMATION IN SUPERHELICAL DOUBLE-STRANDED
CC      DNA. MAY PLAY A ROLE IN MAINTENANCE OF GENOMIC INTEGRITY.
CC      -1- SUBUNIT: COMPONENT OF NUCLEAR RIBOPROTEIN COMPLEXES.
CC      -1- SUBCELLULAR LOCATION: NUCLEAR.
CC      -1- TISSUE SPECIFICITY: UBQUITOUS.
CC      -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG ISOFORM (SHOWN HERE)
CC      AND A SHORT ISOFORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC      -1- DISEASE: A FORM OF MALIGNANT MYXOID LIPOSARCOMA IS CHARACTERIZED
CC      BY A CHROMOSOMAL TRANSLOCATION t(12;16)(q13;p11) THAT INVOLVES FUS
CC      AND CHOP.
CC      -1- DISEASE: A FORM OF ACUTE MYELOID LEUKEMIA (AML) IS CHARACTERIZED
CC      BY A CHROMOSOMAL TRANSLOCATION t(16;21)(p11;q22) THAT INVOLVES
CC      FUS AND ERG.
CC      -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC      -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb.ch/announce/
CC      or send an email to license@isb.sib.ch).
CC      -----
DR      EMBL: S62140; AAB27102.1; -.
DR      EMBL: S62138; AAB27103.1; ALT_SEQ.
DR      EMBL: X71427; CAAS0558.1; ALT_SEQ.
DR      EMBL: X71428; CAAS0559.1; ALT_SEQ.
DR      EMBL: AF071213; AAC35285.1; -.
DR      EMBL: AF071213; AAC35284.1; -.
DR      PIR: S33799; S33799.
DR      HSSP: P09651; 1HA1.
DR      MIM: 137070; -.
DR      MIM: 151900; -.
DR      INTERPRO: IPR000504; -.
DR      INTERPRO: IPR001876; -.
DR      PFMF: PF00076; rrm; 1.
DR      PFMF: PF00641; zf-RanBP; 1.
DR      PROSITE: PS50102; RRM; 1.
DR      PROSITE: PS50030; RRM_RNP_1; FALSE_NEG.
DR      K0 Proto-oncogene; RNA-binding; DNA-binding; Nuclear protein; Repeat;
DR      KW Alternative splicing; Chromosomal translocation; zinc-finger; zinc;
DR      KW Metal-binding.
DR      KW DOMAIN 1 165 GLN/GLY/SER/YTR-RICH.
DR      FT DOMAIN 166 267 GLY-RICH.
DR      FT DOMAIN 285 371 RNA-BINDING (RRM).
DR      FT DOMAIN 371 526 ARG/GLY-RICH.
DR      FT ZN_FING 428 447 C4-TYPE (POTENTIAL).
DR      FT SITE 266 267 BREAKPOINT FOR TRANSLOCATION TO FORM
DR      FT FUS/TLS-CHOP ONCOGENE.
DR      FT VARSPLIC 64 65 FUS/TLS-CHOP (ISOFORM).
DR      FT CONFLICT 338 338 TG -> S (IN SHORT ISOFORM).
DR      FT SEQ 526 AA; 53426 MW; 88C8E263B7905549 CRC64;
OY      Query Match 23.8%; Score 69; DB 1; Length 526;
OY      Best Local Similarity 40.0%; Pred. No. 2;
OY      Matches 18; Conservative 2; Mismatches 21; Indels 4; Gaps 1;
OY      5 YGLRPGSGSQDMWYGLRPGGSS---OHKWSYGLRPGSGSQDMWSYG 45
OY      100 YGQAPPSSTSGSYGSSSSGSSSYGQGPSSGYSQAPSYGQDQSYG 144
DB      100 YGQAPPSSTSGSYGSSSSGSSSYGQGPSSGYSQAPSYGQDQSYG 144
RESULT 12
CONJL_HAPBU STANDARD. PRT. 94 AA.
P51918.0893387.
01-OCT-1996 (rel. 34, Created)

```

DT	30-MAY-2000 (Rel. 39, last sequence update)
DT	30-MAY-2000 (Rel. 39, last annotation update)
DE	GONADOLIBERIN I PRECURSOR (GONADOTROPIN-RELEASING HORMONE I) (GNRH-I)
DE	(LH-RH I) (LULIBERIN I).
GN	GNRH1.
OS	Haplochromis burtoni.
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC	Acanthomorphi; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC	Cichlidae; Astatotilapia.
CC	[1]
RP	SEQUENCE FROM N.A.
RA	MEDLINE=95396797; PubMed=7667296;
RA	White S.A., Kastan T.L., Bond C.T., Adelman J.P., Fernald R.D.:
RT	"Three gonadotropin-releasing hormone genes in one organism suggest
RT	novel roles for an ancient peptide."
RT	Proc. Natl. Acad. Sci. U.S.A. 92:8363-8367(1995).
RN	[2]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=99061842; PubMed=9843638;
RA	White R.B., Fernald R.D.:
RT	"Ontogeny of gonadotropin-releasing hormone (GNRH) gene expression
RT	reveals a distinct origin for GNRH-containing neurons in the
RT	midbrain."
RL	Gen. comp. Endocrinol. 112:322-329(1998).
RN	[3]
RP	SEQUENCE OF 23-32.
RC	TISSUE=PIUTITARY;
RX	MEDLINE=95372591; PubMed=7644702;
RA	Powell J.F.F., Fischer W.H., Park M., Craig A.G., Rivier J.E.,
RA	White S.A., Francis R.C., Fernald R.D., Licht P., Wadby C.,
RA	Sherwood N.M.:
RT	"Primary structure of solitary form of gonadotropin-releasing hormone
RT	(GNRH) in cichlid pituitary; three forms of GNRH in brain of cichlid
RT	and pumpkinseed fish."
RL	Regul. Pept. 57:43-53(1995).
CC	-I- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS. MAY BE
CC	RESPONSIBLE FOR THE REGULATION OF THE HYPOTHALAMIC-PITUITARY-
CC	GONADAL AXIS.
CC	-I- TISSUE SPECIFICITY: SYNTHESIZED IN PREOPTIC NEURONS AND IS
CC	TRANSPORTED TO THE PITUITARY IN THE PREOPTIC-HYPOPHYSAL AXONS.
CC	-I- MASS SPECTROMETRY: MM=1113.9; METHOD=MALDI; RANGE=23-32.
CC	-I- SIMILARITY: BELONGS TO THE GNRH FAMILY.
CC	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL, U01865; AAC59691.1; -
DR	EMBL, AF076961; AAC27716.1; -
DR	INTERPRO: IPR002012; -
DR	PFAM: PF00446; GNRH; 1.
DR	PROSITE: PS00473; GNRH; 1.
KM	Cleavage on pair of basic residues; Hormone; Amidation; Hypothalamus;
KM	Signal; Multigene family.
FT	SIGNAL
FT	CHAIN
FT	PEPTIDE
FT	PEPTIDE
FT	MOD_RES
FT	MOD_RES
FT	CONFLICT
FT	SEQUENCE
FT	94 AA: 10362 MW; E57DDBA3333278D7 CRC64;
FT	94
FT	23
FT	36
FT	23
FT	32
FT	94
FT	86
FT	1
FT	22
FT	94
FT	32
FT	94
FT	23
FT	32
FT	94
FT	86
FT	1
FT	22
FT	94
FT	32
FT	94
FT	23
FT	32
FT	94
FT	86
FT	1
FT	22
FT	94
FT	32
FT	94
FT	23
FT	32
FT	94
FT	86
FT	1
FT	22
FT	94
FT	32
FT	94
FT	23
FT	32
FT	94
FT	86
FT	1
FT	22
FT	94
FT	32
FT	94
FT	23
FT	32
FT	94
FT	86
FT	1
FT	22
FT	94
FT	32
FT	94
FT	23
FT	32
FT	94
FT	86
FT	1
FT	22
FT	94
FT	32
FT	94
FT	23
FT	32
FT	94
FT	86
FT	1
FT	22
FT	94
FT	32
FT	94
FT	23
FT	32
FT	94
FT	86
FT	1
FT	22
FT	94
FT	32
FT	94
FT	23
FT	32
FT	94
FT	86
FT	1
FT	22
FT	94
FT	32
FT	94
FT	23
FT	32
FT	94
FT	86
FT	1
FT	22
FT	94
FT	32
FT	94
FT	23
FT	32
FT	94
FT	86
FT	1
FT	22
FT	94
FT	32
FT	94
FT	23
FT	32
FT	94
FT	

Query Match	23.6%;	Score 68.5;	DB 1;	Length 94;
Best Local Similarity	65.0%;	Pred. No. 0.42;		
Matches	13;	Conservative	1;	Mismatches 5;
			Indels	1;
			Gaps	1

Search completed: March 2, 2001, 11:01:21
Job time: 436 sec

RC STRAIN-JD210;
RX MEDLINE-91064385; Pubmed-2248978;
RA Chen J.H., Chen Z.Y., Chow T.Y., Chen J.C., Tan S.T., Hsu W.H.;
RT "Nucleotide sequence and expression of the isoamylase gene from an
isoamylase-hyperproducing mutant, *Pseudomonas amyloidermosa* JD210.";
RL Biochim. Biophys. Acta 1087:309-315(1990).
RN [3]
RP SEQUENCE OF 744-776 FROM N.A.
RC STRAIN-SB-15;
RX MEDLINE-89327147; Pubmed-2753857;
RA Amemura A., Fujita M., Futai M.;
RT "Transcription of the isoamylase gene (iam) in *Pseudomonas*
RT *amyloidermosa* SB-15.";
RL J. Bacteriol. 171:4320-4325(1989).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE-98387895; Pubmed-9719642;
RA Katsuya Y., Mezaki Y., Kubota M., Matsuura Y.;
RT "Three-dimensional structure of *Pseudomonas* isoamylase at 2.2-A
RT resolution.";
RL J. Mol. Biol. 281:885-897(1998).
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,6-ALPHA-D-GLUCOSIDIC BRANCH
CC LINKAGES IN GLYCOGEN, AMYLOPECTIN AND THEIR BETA-LIMITS DEXTRINS.
CC -1- INDUCTION: BY MALTOSE.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY, ISOAMYLASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03871; AAA25854.1; -;
DR EMBL; X13378; CAA31754.1; -;
DR PIR; A28109; A28109.
DR PDB; 1BF2; 12-AUG-98.
DR INTERPRO; IPR000461; -;
DR PRAM; PF00128; alpha-amyase; 1.
KW Hydrolyase; Glycosidase; Signal; 3D-structure.
FT SIGNAL 1 26
FT CHAIN 27 776
FT ACT_SITE 401 401
FT ACT_SITE 481 481
FT ACT_SITE 536 536
FT DISULFID 410 422
FT DISULFID 546 616
FT DISULFID 738 766
FT CONFLICT 8 8
FT CONFLICT 126 126
FT CONFLICT 169 171
FT CONFLICT 386 386
FT CONFLICT 413 416
FT CONFLICT 454 489
FT CONFLICT 555 556
FT CONFLICT 650 657
SQ SEQUENCE 776 AA; 83626 MW; F73BBF8040246169 CRC64;
A -> G (IN REF. 1).
F -> C (IN REF. 1).
GAS -> AH (IN REF. 1).
L -> V (IN REF. 1).
GAYT -> AVH (IN REF. 1).
SGLDLFAEPWAIIGNSYOLGFGPGQWSENGLFRDS ->
TWICLRNLGSPSAITRTSWVDSRRVAVENSVPRQ (IN
REF. 1).
WP -> S (IN REF. 1).
AFRAHRA -> RSARHIP (IN REF. 1).
Query Match 22.88; Score 66; DB 1; Length 776;
Best Local Similarity 35.14; Pred. NO. 5.9;
Matches 13; Conservative 6; Mismatches 16; Indels 2; Gaps 1;
OY 5 YGLRPGSGSDMWYGLRPGSSOH--WSYGLRPGSGS 39
DB 544 YSCNANNSQAMPYGPSPDGGTSTNYSWDQMSAGTGA 580

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:55:41 ; Search time 56.87 Seconds
(Without alignments)
58.504 Million cell updates/sec

Title: US-09-306-689-11

Perfect score: 290
Sequence: 1 OHMSYGLRPGSSQDMSTYGL.....SYGLRPGSSQDMSTYGLRPG 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR.66:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	77	26.6	445	2 A27263	50K spicule matrix
2	76.5	26.4	67	2 I78541	gonadoliberin prec
3	76.5	26.4	92	1 RHHUG	gonadoliberin prec
4	74	25.5	90	1 RHMSG	gonadoliberin prec
5	74	25.5	92	1 RHMSG	gonadoliberin prec
6	72	24.8	528	2 G02127	tus-like protein
7	72	24.8	1068	2 T42382	gonadoliberin prec
8	71	24.5	414	2 A36443	seminal vesicle se
9	70.5	24.3	171	2 T16933	hypothetical prote
10	69	23.8	462	4 S33798	FUS/CHOP mutant fu
11	69	23.8	526	1 S33799	RNA-binding protei
12	69	23.8	2055	2 T31617	gonadoliberin prec
13	68.5	23.6	98	2 T50739	gonadoliberin prec
14	68.5	23.6	705	2 T00975	hypothetical prote
15	68	23.4	335	2 A43900	spicule matrix pro
16	66.5	22.9	89	2 T51423	gonadoliberin prec
17	66.5	22.9	313	2 T04776	hypothetical prote
18	66.5	22.9	345	2 A58519	hypothetical 345 p
19	66.5	22.9	772	2 T02098	phosphatidylinosit
20	66	22.8	776	2 S13470	isoamylase (EC 3.2
21	66	22.8	776	2 A37035	isoamylase (EC 3.2
22	66	22.8	798	2 T25104	hypothetical prote
23	63	21.7	10	1 RHMSG	gonadoliberin - pi
24	63	21.7	10	1 RHMSG	gonadoliberin - pi
25	63	21.7	646	2 T43677	ATP-dependent RNA
26	63	21.7	2183	2 T37218	hypothetical prote
27	62.5	21.6	343	2 T05221	hypothetical prote
28	62	21.4	190	2 T29068	hypothetical prote
29	62	21.4	848	2 S02262	glutenin high mole

30	61.5	21.2	465	2 F75524	hypothetical prote
31	61	21.0	118	2 T32631	hypothetical prote
32	61	21.0	538	2 T49418	hypothetical prote
33	61	21.0	767	2 S47639	acetylcholinestera
34	60.5	20.9	314	2 T08675	hypothetical prote
35	60	20.7	92	2 I50644	gonadoliberin I pr
36	60	20.7	167	2 S21359	keratin, type I, c
37	60	20.7	481	2 T15172	hypothetical prote
38	60	20.7	505	2 A39127	sucrose porin scry
39	60	20.7	505	2 S15193	sucrose porin scry
40	60	20.7	1653	2 G65028	hypothetical prote
41	60	20.7	2206	2 JC5280	voltage-dependent
42	59.5	20.5	683	2 T51821	1-phosphatidylinos
43	59.5	20.5	769	2 T22256	hypothetical prote
44	59.5	20.5	838	1 ESWTHW	glutenin, high mol
45	59.5	20.5	990	2 T38274	probable transcrip

ALIGNMENTS

```

RESULT 1
A27263
50K spicule matrix protein precursor - sea urchin (Strongylocentrotus purpuratus)
C:Species: Strongylocentrotus purpuratus (purple urchin)
C>Date: 19-Nov-1988 #sequence_revision 12-May-1994 #text_change 17-Mar-1999
C:Accession: A27263; A49766
R:Sucov, H.M.; Benson, S.; Robinson, J.J.; Britten, R.J.; Wilt, F.; Davidson, E.H.
Dev. Biol. 120, 507-519, 1987
A:Title: A lineage-specific gene encoding a major matrix protein of the sea urchin em
A:Reference number: A27263; MUID:87162883
A:Accession: A27263
A:Molecule type: mRNA
A:Residues: 1-199, 'GTSLSANRAMPKALVINOAGVGDNOALVINOEWGDNQALVINOEWGDNQAVINPVWVGDN
AGDNQWEGDNQALVINQWVGDNOAWASNRIRIRTRTRTRITRITRPGSTDPVCFKRQTRHNLSSD' <SUC>
A:Note: this sequence has been corrected in reference A49766
R:Katoh-Fukui, Y.; Noce, T.; Ueda, T.; Fujiwara, Y.; Hashimoto, N.; Higashinakagawa,
Dev. Biol. 145, 201-202, 1991
A:Title: The corrected structure of the SMO50 spicule matrix protein of Strongylocentr
A:Reference number: A49766; MUID:91209669
A:Accession: A49766
A:Molecule type: mRNA
A:Residues: 1-445 <KAT>
A:Cross-references: GB:M16231; NID:g161560; PID:g161562; GB:M15760
C:Gene: SMO50
A:Gene: SMO50
A:Introns: 35/72
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-445/Product: 50K spicule matrix protein #status predicted <MAT>
F:212-307,321-411/Region: 13-residue repeats (Q-P-G-F/M-W-G-N/G-Q-P-G-V/M-G-G-R)
F:411-430/Region: 3-(or 2)-residue repeats (P-N-N or P-N)

Query Match 26.6% Score 77; DB 2; Length 445;
Best Local Similarity 38.6% Pred. No. 0.47;
Matches 17; Conservative 7; Mismatches 20; Indels 0; Gaps 0;

QY 6 GLRPGSSQDMSTYGLRPGSSQDMSTYGLRPGSSQDMSTYGLRPG 49
DB 237 GRPFGNGQPMGGRQPMGNGQPMGNGQPMGNGQPMGNGQPMG 280

RESULT 2
I78541
gonadoliberin precursor - rhesus macaque (fragment)
N:Alternate names: luteinizing hormone releasing hormone
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C:Accession: I78541
R:Ma, Y.J.; Costa, M.E.; Ojeda, S.R.
Neuroendocrinology 60, 346-359, 1994
A:Title: Developmental expression of the genes encoding transforming growth factor al
A:Reference number: I58134; MUID:95124501

```

A:Accession: I78541
A>Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: mRNA
A:Residues: 1-67 <RES>
A:Cross-references: GB:S75918; NID:g912831; PIDN:AAB33096.1; PID:g912832
C:Superfamily: gonadoliberin

Query Match 26.4%; Score 76.5; DB 2; Length 67;
Best Local Similarity 78.9%; Pred. No. 0.077;
Matches 15; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Oy 23 GGSSQHWSTGRLPGSGSOD 41
 | |||||
Db 2 GCSSQHWSTGLRFG-GKRD 19

RESULT 3
RHNG

gonadoliberin precursor - human
N:Alternate names: gonadotropin releasing hormone (GNRH); luteinizing hormone releasing
N:Contains: gonadoliberin-associated protein (GAP); progadoliberin
C:Species: Homo sapiens (man)
C>Date: 17-Mar-1987 #sequence.revision 21-Jul-1995 #text.change 18-Jun-1999
C:Accession: S05308; A26173; A93342; A90108; A01410; S45718
R:Hayflick, J.S.; Adelman, J.P.; Seeburg, P.H.
Nucleic Acids Res. 17, 6403-6404, 1939
A>Title: The complete nucleotide sequence of the human gonadotropin-releasing hormone gene
A:Reference number: S05308; MUID:89356682
A:Accession: S05308
A>Status: translation not shown

A:Molecule type: DNA
A:Residues: 1-92 <HAY>
A:Cross-references: EMBL:X15215; NID:g31955; PIDN:CAA33285.1; PID:g31956
R:Adelman, J.P.; Mason, A.J.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 173-183, 1986
A>Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotro
A:Reference number: A94090; MUID:86094338
A:Accession: A26173

A:Molecule type: mRNA
A:Residues: 1-92 <ADE>
A:Cross-references: GB:M12578; NID:g183418; PIDN:AAA5916.1; PID:g386749
A:Experimental source: Hypothalamus
R:Seeburg, P.H.; Adelman, J.P.
Nature 311, 666-668, 1984
A>Title: Characterization of cDNA for precursor of human luteinizing hormone releasing h
A:Reference number: A93342; MUID:85012739
A:Accession: A93342

A:Molecule type: mRNA
A:Residues: 1-15, 'S', 17-92 <SEP>
A:Cross-references: GB:X01059; NID:g34356; PIDN:CAA25526.1; PID:g34357
A:Experimental source: Placenta
R:Tan, L.; Rousseau, P.
Biochem. Biophys. Res. Commun. 109, 1061-1071, 1982
A>Title: The chemical identity of the immunoreactive LHRH-like peptide biosynthesized in
A:Reference number: A90108; MUID:83126573
A:Accession: A90108

A:Molecule type: protein
A:Residues: 24-33 <TAN>
A:Experimental source: Placental trophoblasts
R:Ishibashi, D.; Koch, Y.; Pitzer, F.; Fridkin, M.; Dantes, A.; Baumeister, W.; Amsterda
FEBS Lett. 346, 203-206, 1994
A>Title: Sequential degradation of the neuropeptide gonadotropin-releasing hormone by th
A:Reference number: S45718; MUID:94233597
A:Contents: annotation; degradation pathway of synthetic hormone
C:Genetics:

A:Gene: GDB:GNRH; LHRH; GRH
A:Cross-references: GDB:I33746; OMIM:227200; OMIM:152760
A:Map position: Bp21-8p11.2
A:Introns: 47/3; 79/3

C:Function:
A:Note: gonadoliberin stimulates pituitary secretion of lutropin and follitropin
A:Note: gonadoliberin-associated protein may have prolactin release inhibiting activity

```

C:Superfamily: gonadoliblerin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-92/Product: progadoliblerin #status predicted <PGN>
F:24-33/Product: gonadoliblerin #status experimental <MAT>
F:37-92/Product: gonadoliblerin-associated protein #status predicted <GAP>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match      26.4%; Score 76.5; DB 1; Length 92;
Best Local Similarity 78.9%; Pred. No. 0.11;
Matches 15; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

OY 23 GGSSQHWSTGLRPGSGD 41
    | | | | | | | | | | | | | | | | | |
Db 20 GGSSQHWSTGLRPG-GKRD 37

RESULT 4
RMSG
gonadoliblerin precursor - mouse
N:Alternate names: gonadotropin-releasing hormone (GNRH); luteinizing hormone releasing
N:Contains: gonadoliblerin; gonadoliblerin-associated protein (GAP)
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1993 #sequence_revision 18-Mar-1997 #text_change 18-Jun-1999
C:Accession: A47578
R:Maeson, A.J.; Hayflick, J.S.; Zoeller, R.T.; Young III, W.S.; Phillips, H.S.; Nikolai
Science 234, 1366-1371, 1986
A:Title: A deletion truncating the gonadotropin-releasing hormone gene is responsible
A:Reference number: A47578; MUID:87065928
A:Accession: A47578
A:Molecule type: DNA
A:Residues: 1-90 <MAS>
A:Cross-references: EMBL:M14872; NID:q193576; PIDN:AAA37717.1; PID:9387175
C:Genetics:
A:Introns: 45/3; 77/3
C:Function:
A:Description: gonadoliblerin stimulates pituitary secretion of lutropro and follitropo
A:Note: gonadoliblerin-associated protein may have prolactin release inhibiting activi
C:Superfamily: gonadoliblerin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid
F:1-23/Domain: signal sequence #status predicted <SIG>
F:22-31/Product: gonadoliblerin #status predicted <GLB>
F:35-90/Product: gonadoliblerin-associated protein #status predicted <GAP>
F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predia
F:31/Modified site: amidated carboxyl end (Gly) (amide in mature form from following

Query Match      25.5%; Score 74; DB 1; Length 90;
Best Local Similarity 92.9%; Pred. No. 0.19;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 23 GGSSQHWSTGLRPG 36
    | | | | | | | | | | | | | | | | | |
Db 18 GGSSQHWSTGLRPG 31

RESULT 5
RHRG
gonadoliblerin precursor - rat
N:Alternate names: gonadoliblerin-associated protein (GAP); gonadotropin releasing hor
N:Contains: gonadoliblerin; prolactin release-inhibiting factor
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1999
C:Accession: A40147; B26173; A48410
R:Bond, C.T.; Hayflick, J.S.; Seeburg, P.H.; Adelman, J.P.
Mol. Endocrinol. 3, 1257-1262, 1989
A:Title: The rat gonadotropin-releasing hormone: SH locus: structure and hypothalamic
A:Reference number: A40147; MUID:89384661
A:Accession: A40147
A:Molecule type: DNA
A:Residues: 1-92 <BON>

```

```

A:CROSS-references: GB:M1670; NID:g204447; PIDN:AAA1264.1; PID:g204448
R:Adelman, J.P.; Mason, A.U.; Hayflick, J.S.; Seeburg, P.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 179-183, 1986
A>Title: Isolation of the gene and hypothalamic cDNA for the common precursor of gonadotropin releasing hormone
A:Reference number: A94090; MUID:86094338
A:Accession: B26173
A:Molecule type: mRNA
A:Residues: 1-92 <ADE>
A:CROSS-references: GB:M12579; NID:g204445; PIDN:AAA1263.1; PID:g204446
R:Maier, C.C.; Marchetti, B.; LeBoeuf, R.D.; Blalock, J.E.
Cell. Mol. Neurobiol. 12, 447-454, 1992
A>Title: Thyrocytes express a RNA that is identical to hypothalamic luteinizing hormone releasing hormone
A:Reference number: A48410; MUID:93105480
A:Accession: A48410
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <MAI>
A:CROSS-references: GB:S50870; NID:g262059; PIDN:AAB24572.1; PID:g262060
A:Experimental source: thymus
A>Note: sequence extracted from NCBI backbone (NCBIN:121082, NCBIP:121083)
C:Genetics:
A:Introns: 47/3; 79/3
C:Function:
A>Description: stimulates pituitary secretion of lutropin and follitropin
A>Note: gonadoliberin associated protein may have prolactin release inhibiting activity
C:Superfamily: gonadoliberin
C:Keywords: amidated carboxyl end; hormone; hypothalamus; placenta; pyroglyutamic acid; n
F:1-23/Domin: signal sequence #status predicted <SIG>
F:24-92/Product: progonaodoliberin #status predicted <PCN>
F:24-33/Product: gonadoliberin #status predicted <GLN>
F:37-92/Product: prolactin release-inhibiting factor #status predicted <PIF>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:33/Modified site: amidated carboxyl end (Gly) (amide in mature form from following glyx

```

guanylate cyclase (EC 4.6.1.2) 1, retinal - chicken (fragment)
N:Alternate names: guanylyl cyclase; retinal guanylate cyclase 1
C:Species: Gallus gallus (chicken)
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jun-2000
C:Accession: T42382
R:Sample-Rowland, S.L.; Lee, N.R.; Van Hooser, J.P.; Palczewski, K.; Baehr, W.
A:Title: A null mutation in the photoreceptor guanylate cyclase gene causes the retin
A:Reference number: 722141; MUID:98115910
A:Accession: T42382
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1068 <SEM>
A:Cross-references: EMBL:AF036942; NID:g2828021; PID:g2828022; PID:AA04500.1
C:Genetics:
A:Gene: GCL
C:Superfamily: membrane-bound guanylate cyclase; guanylate cyclase catalytic domain h
C:Keywords: cGMP biosynthesis; glycoprotein; phosphorus-oxygen lyase; transmembrane p

Query Match 24.8%; Score 71; DB 2; Length 1068;
Best Local Similarity 39.0%; Pred. No. 3.8;
Matches 16; Conservative 3; Mismatches 14; Indels 8; Gaps 2;

QY 10 GSGSQDWS-YGLRPG-----GSSQHWYGLRPGSGSQDW 42
Db 356 GKGDRLMPVYGLPEGRGLAYRGSHVHPHSSSGTDSGCW 396
| : | | | | | : | | | : | : | |
RESULT 8
A36443
seminal vesicle secretory protein II precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 08-Mar-1991 #sequence_revision 08-Mar-1991 #text_change 05-Nov-1999
C:Accession: A36443
R:Harris, S.E.; Harris, M.A.; Johnson, C.M.; Bean, M.F.; Dodd, J.G.; Matusik, R.J.; C
J. Biol. Chem. 265, 9896-9903, 1990
A:Title: Structural characterization of the rat seminal vesicle secretion II protein
A:Reference number: A36443; MUID:90277684
A:Accession: A36443
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-414 <HAR>
A:Cross-references: GB:J05443; NID:g207114; PID:AAA42192.1; PID:g207115

Query Match 24.5%; Score 71; DB 2; Length 414;
Best Local Similarity 42.9%; Pred. No. 1.8;
Matches 18; Conservative 3; Mismatches 21; Indels 0; Gaps 0;

QY 4 SYGLRPGSGSQDWSYGLRPGSGQHWMSYGLRPGSGSQDWSYG 45
Db 184 SFGQMKSSGSGQVSKFGQMKASESQIKSFQKSSGQGLSYTG 225
| : | | | | | : | | | : | | | | |
RESULT 9
T16933
hypothetical protein T24D8.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16933
R:Martin, J.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid T24D8.
A:Reference number: Z18609
A:Accession: T16933
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-171 <MAR>
A:Cross-references: EMBL:U04042; NID:g1065529; PID:g1065532; PID:AAA81446.1; CESP:IT2
A:Genetics:
A:Gene: CESP:T24D8.5

THIS PAGE BLANK (08PT0)

us-09-306-689-10.rap

/cgn2_6/
/cgn2_6/
/cgn2_6/

/cgn2_6/

/cgn2_6/ptodata/2/paa/US093_COMB.pcp:US-09-374-721a-31 + 105.00	236.28	0.0004	2
/cgn2_6/ptodata/2/paa/US07_COMB.pcp:US-07-761-849-11 + 105.00	235.87	0.0004	21
/cgn2_6/ptodata/2/paa/US07_COMB.pcp:US-07-761-849-12 + 105.00	235.87	0.0004	21
/cgn2_6/ptodata/2/paa/US084_COMB.pcp:US-08-476-013-11 + 105.00	235.87	0.0004	211
/cgn2_6/ptodata/2/paa/US084_COMB.pcp:US-08-476-013-12 + 105.00	235.87	0.0004	211
seq_name: /cgn2_6/ptodata/2/paa/US090_COMB.pcp:US-09-019-010-4			
seq_documentation_block:			
: Sequence 4, Application US/09019010			
: GENERAL INFORMATION:			
: APPLICANT: HARLAND, RICHARD			
: APPLICANT: MANNS, JOHN G.			
: APPLICANT: ACRES, STEPHEN D.			
: TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS			
: MOLECULES			
: NUMBER OF SEQUENCES: 6			
: CORRESPONDENCE ADDRESS:			
: ADDRESSEE: ROBINS & ASSOCIATES			
: STREET: 90 MIDDLEFIELD ROAD, SUITE 200			

```

seq_documentation_block:
Sequence 4, Application US/09019010
GENERAL INFORMATION:
APPLICANT: HARLAND, RICHARD
APPLICANT: MANN'S, JOHN G.
APPLICANT: ACRES, STEPHEN D.
TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROBINS & ASSOCIATES
STREET: 90 MIDDLEFIELD ROAD, SUITE 200
CITY: MENLO PARK
STATE: CA
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

seq_documentation_block:
  Sequence 4, Application US/09019010
  GENERAL INFORMATION:
    APPLICANT: HARLAND, RICHARD
    APPLICANT: MANN, JOHN G.
    APPLICANT: ACRES, STEPHEN D.
    TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS
    TITLE OF INVENTION: MOLECULES
    NUMBER OF SEQUENCES: 6
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: ROBINS & ASSOCIATES
      STREET: 90 MIDDLEFIELD ROAD, SUITE 200
      CITY: MENLO PARK
      STATE: CA
      COUNTRY: USA
      ZIP: 94025
    COMPUTER READABLE FORM:
      MEDIUM TYPE: floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/09/019,010
      FILING DATE: 05-FEB-1998

```

```

seq_documentation_block:
Sequence 4, Application US/09019010
GENERAL INFORMATION:
APPLICANT: HARLAND, RICHARD
APPLICANT: MANN, JOHN G.
APPLICANT: ACRES, STEPHEN D.
TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS
NUMBER OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROBINS & ASSOCIATES
STREET: 90 MIDDLEFIELD ROAD, SUITE 200
CITY: MENLO PARK
STATE: CA
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/019,010
FILING DATE: 05-FEB-1998
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/036,883
FILING DATE: 05-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0035
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 325-7812
TELEFAX: (650) 325-7823
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-019-010-4

alignment_scores:
Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-09-019-010-4 ..
Align seg 1/1 to: US-09-019-010-4 from: 1 to: 49

1 CAGCATTTGAGCTAGCGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
1 GlnHsTrpSerTyrGlyLeuArgProGlySerGlnAspTrpSe 17
17 TylctyLeuArgProGlySerGlnHsTrpSerTyrGlyLeuA 34
51 GCCTGGCAGCGGATGCCAGATTGGAGCTACGGCCTCGGCT 147
|||||
34 rgrProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49

```

seq_name: /cgn2_6/ptodata/2/paa/US092_COMB.pep:US-09-249-447A-7

seq_documentation_block:

```
; Sequence 7, Application US/09249447A
; GENERAL INFORMATION:
; APPLICANT: Manns, Jack G.
; TITLE OF INVENTION: PASSIVE IMMUNIZATION AS A TREATMENT FOR HORMONE
; FILE REFERENCE: 9001-0045
; CURRENT APPLICATION NUMBER: US/09/249,447A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/075,637
; PRIOR FILING DATE: 1998-02-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 7
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GnRH-2, Fig.
; US-09-249-447A-7
```

alignment_scores:

Quality:	290.00	Length:	49
Ratio:	5.918	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-306-689-10 x US-09-249-447A-7 ..

Align seg 1/1 to: US-09-249-447A-7 from: 1 to: 49

```
1 CAGCATGGAGCTACGGCTGCGCCCTGCGAGCGGTTCTCAAGATTGGAG 50
|||||
1 GlnHstPserTyrGlyLeuArgProGlySerGlnHstPserTyrGlyLeuA 17
|||||
51 CTAGGGCTGCGTCCGGGTGCTCTAGCAGCATTTGAGCTACGGCTGC 100
|||||
17 rTyrGlyLeuArgProGlySerGlnHstPserTyrGlyLeuA 34
|||||
101 GCCCTGGAGCGGTAGCCAGATTGGAGCTACGGCTGCGGGT 147
|||||
34 rGProGlySerGlySerGlnAspTyrGlyLeuArgProGly 49
```

seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:US-09-305-924-11

seq_documentation_block:

```
; Sequence 11, Application US/09305924A
; GENERAL INFORMATION:
; APPLICANT: Jack G. Manns
; APPLICANT: Stephen D. Acres
; APPLICANT: Richard Harland
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305,924A
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,217
; EARLIER FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 11
; LENGTH: 49
; TYPE: PRT
; ORGANISM: GnRH
; US-09-305-924-11
```

alignment_scores:

Quality:	290.00	Length:	49
----------	--------	---------	----

Ratio:	5.918	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-306-689-10 x US-09-305-924-11 ..

Align seg 1/1 to: US-09-305-924-11 from: 1 to: 49

```
1 CAGCATGGAGCTACGGCTGCGCCCTGCGAGCGGTTCTCAAGATTGGAG 50
|||||
1 GlnHstPserTyrGlyLeuArgProGlySerGlnHstPserTyrGlyLeuA 17
|||||
51 CTAGGGCTGCGTCCGGGTGCTCTAGCAGCATTTGAGCTACGGCTGC 100
|||||
17 rTyrGlyLeuArgProGlySerGlnHstPserTyrGlyLeuA 34
|||||
101 GCCCTGGAGCGGTAGCCAGATTGGAGCTACGGCTGCGGGT 147
|||||
34 rGProGlySerGlySerGlnAspTyrGlyLeuArgProGly 49
```

seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:US-09-306-689-11

seq_documentation_block:

```
; Sequence 11, Application US/09306689B
; GENERAL INFORMATION:
; APPLICANT: Robbins, Sarah C.
; TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
; FILE REFERENCE: 9001-0047
; CURRENT APPLICATION NUMBER: US/09/306,689B
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: US 60/088,024
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 11
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; US-09-306-689-11
```

alignment_scores:

Quality:	290.00	Length:	49
Ratio:	5.918	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

US-09-306-689-10 x US-09-306-689-11 ..

Align seg 1/1 to: US-09-306-689-11 from: 1 to: 49

```
1 CAGCATGGAGCTACGGCTGCGCCCTGCGAGCGGTTCTCAAGATTGGAG 50
|||||
1 GlnHstPserTyrGlyLeuArgProGlySerGlnHstPserTyrGlyLeuA 17
|||||
51 CTAGGGCTGCGTCCGGGTGCTCTAGCAGCATTTGAGCTACGGCTGC 100
|||||
17 rTyrGlyLeuArgProGlySerGlnHstPserTyrGlyLeuA 34
|||||
101 GCCCTGGAGCGGTAGCCAGATTGGAGCTACGGCTGCGGGT 147
|||||
34 rGProGlySerGlySerGlnAspTyrGlyLeuArgProGly 49
```

seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:US-09-383-912-4

seq_documentation_block:

```
; Sequence 4, Application US/09383912
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
```

```

; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-383-912-4

alignment_scores:
  Quality: 290.00      Length: 49
  Ratio: 5.918         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-09-383-912-4 ..
Align seg 1/1 to: US-09-383-912-4 from: 1 to: 49

1 CAGCATTGGAGCTAGCGCCTGCGCCCTGCGAGCGGTTCACAGATTGGAG 50
|||||
1 GlnHstIpSerTyrGlyLeuArGProGlySerGlnAspTyrSe 17
17 rTyrGlyLeuArGProGlyGlySerSerGlnHstIpSerTyrGlyLeuA 34
51 CTACGGCCTGCGTCCGGGTGCTCTAGCCAGCATTGGAGCTAGCGCCTGC 100
|||||
17 rTyrGlyLeuArGProGlyGlySerSerGlnHstIpSerTyrGlyLeuA 34
101 GCCCTGGAGCGGTAGCCAGATTGGAGCTAGCGCCTGCGGT 147
|||||
34 rGProGlySerGlySerGlnAspTyrSerTyrGlyLeuArGProGly 49

seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:US-09-383-912-10

seq_documentation_block:
; Sequence 10, Application US/09383912
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
```

```

; STATE: CA
; COUNTRY: USA
; ZIP: 94301
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-383-912-10

alignment_scores:
  Quality: 290.00      Length: 49
  Ratio: 5.918         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x US-09-383-912-10 ..
Align seg 1/1 to: US-09-383-912-10 from: 1 to: 544

1 CAGCATTGGAGCTAGCGCCTGCGCCCTGCGAGCGGTTCACAGATTGGAG 50
|||||
494 GlnHstIpSerTyrGlyLeuArGProGlySerGlnAspTyrSe 510
510 rTyrGlyLeuArGProGlyGlySerSerGlnHstIpSerTyrGlyLeuA 527
101 GCCCTGGAGCGGTAGCCAGATTGGAGCTAGCGCCTGCGGT 147
|||||
527 rGProGlySerGlySerGlnAspTyrSerTyrGlyLeuArGProGly 542

seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:US-09-305-924-13

seq_documentation_block:
; Sequence 13, Application US/09305924A
; GENERAL INFORMATION:
; APPLICANT: Jack G. Manns
; APPLICANT: Stephen D. Acres
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305,924A
; EARLIER FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,217
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 13
; LENGTH: 695
; TYPE: PRT
; ORGANISM: GNRH
```

US-09-305-924-13

alignment_scores:

Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-306-689-10 x US-09-305-924-13 ..

Align seg 1/1 to: US-09-305-924-13 from: 1 to: 695

```

1 CAGCATTTGAGCTACGGCTGCGCCCTGCGACGGCTTCTCAAGATTGGAG 50
|||||
9 GlnHistrpserTyrGlyLeuArgProGlySerGlnHistrpse 25
51 CTACGGCTGCTGCTCCGGGTGCTCTACGCAGCATTTGAGCTGAGCGCTGC 100
|||||
25 rTyrGlyLeuArgProGlySerSerGlnHistrpserTyrGlyLeuA 42
101 GCCCTGGCAGCGGTAGCCAGATTGGAGCTACGGCTGCGCGGT 147
|||||
42 rGProGlySerGlySerGlnAspTrpserTyrGlyLeuArgProGly 57

```

seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:US-09-306-689-13

seq_documentation_block:

```

; Sequence 13, Application US/09306689B
; GENERAL INFORMATION:
; APPLICANT: Robbins, Sarah C.
; TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
; FILE REFERENCE: 9001-0047
; CURRENT APPLICATION NUMBER: US/09/306,689B
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: US 60/088,024
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 695
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; OTHER INFORMATION: defines a leukotoxin gnRH chimera
US-09-306-689-13

```

alignment_scores:

Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-306-689-10 x US-09-306-689-13 ..

Align seg 1/1 to: US-09-306-689-13 from: 1 to: 695

```

1 CAGCATTTGAGCTACGGCTGCGCCCTGCGACGGCTTCTCAAGATTGGAG 50
|||||
9 GlnHistrpserTyrGlyLeuArgProGlySerGlnHistrpse 25
51 CTACGGCTGCTGCTCCGGGTGCTCTACGCAGCATTTGAGCTGAGCGCTGC 100
|||||
25 rTyrGlyLeuArgProGlySerSerGlnHistrpserTyrGlyLeuA 42
101 GCCCTGGCAGCGGTAGCCAGATTGGAGCTACGGCTGCGCGGT 147
|||||
42 rGProGlySerGlySerGlnAspTrpserTyrGlyLeuArgProGly 57

```

seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:US-09-383-912-16

seq_documentation_block:

; Sequence 16, Application US/09383912

```

; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSER: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-383-912-16

```

alignment_scores:

Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-306-689-10 x US-09-383-912-16 ..

Align seg 1/1 to: US-09-383-912-16 from: 1 to: 699

```

1 CAGCATTTGAGCTACGGCTGCGCCCTGCGACGGCTTCTCAAGATTGGAG 50
|||||
9 GlnHistrpserTyrGlyLeuArgProGlySerGlnHistrpse 25
51 CTACGGCTGCTGCTCCGGGTGCTCTACGCAGCATTTGAGCTGAGCGCTGC 100
|||||
25 rTyrGlyLeuArgProGlySerSerGlnHistrpserTyrGlyLeuA 42
101 GCCCTGGCAGCGGTAGCCAGATTGGAGCTACGGCTGCGCGGT 147
|||||
42 rGProGlySerGlySerGlnAspTrpserTyrGlyLeuArgProGly 57

```

seq_name: /cgn2_6/ptodata/2/paa/US093_COMB.pep:US-09-383-912-8

seq_documentation_block:

```

; Sequence 8, Application US/09383912
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; APPLICANT: MANN, JOHN G.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS

```

```

: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: REED & ROBINS LLP
: STREET: 285 HAMILTON AVENUE, SUITE 200
: CITY: PALO ALTO
: STATE: CA
: COUNTRY: USA
: ZIP: 94301
: COMPUTER READABLE FORM:
: MEDIUM TYPE: FLOPPY disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/383,912
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/694,865
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: MCCracken, THOMAS P.
: REGISTRATION NUMBER: 38,548
: REFERENCE/DOCKET NUMBER: 9001-0016.22
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415)327-3400
: TELEFAX: (415)327-3231
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 977 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-383-912-8

```

```

alignment_scores:
  Quality: 290.00      Length: 49
  Ratio: 5.918        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

```

alignment_block:

US-09-306-689-10 x US-09-383-912-8 ..

Align seg 1/1 to: US-09-383-912-8 from: 1 to: 977

```

1 CAGCATGTGAGCTACGGCCCTGGCCCTGGCAGCGGTTCTCAAGATTGAG 50
|||||
927 GlnHSTPserTyrGlyLeuArpProGlySerGlnAspTyrPse 943
51 CTACGGCCTGCGTCCGGGTGGCTCTAGCCAGCATTTGGAGCTACGGCTTC 100
|||||
943 rTyGlyLeuArpProGlySerGlnHSTPserTyrGlyLeuA 960
101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCTGCGCGGT 147
|||||
960 rGProGlySerGlySerGlnAspTyrPserTyrGlyLeuArpProGly 975

seq_name: /cgn2_6/plodata/2/paa/PCTUS_COMB.pep:PCT-US00-22121-35
seq_documentation_block:
: Sequence 35, Application PC/TUS0022121
: GENERAL INFORMATION:
: APPLICANT: Proteinix Company
: TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
: FILE REFERENCE: IGN-2004WO
: CURRENT APPLICATION NUMBER: PCT/US00/22121
: CURRENT FILING DATE: 2000-08-14
: PRIOR APPLICATION NUMBER: 09/026,276
: PRIOR FILING DATE: 1998-02-19
: PRIOR APPLICATION NUMBER: US 09/374,721
: PRIOR FILING DATE: 1999-08-13
: NUMBER OF SEQ ID NOS: 35

```

```

: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 35
: LENGTH: 40
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: polypeptide
: OTHER INFORMATION: antigen
PCT-US00-22121-35

```

```

alignment_scores:
  Quality: 199.50      Length: 49
  Ratio: 5.250        Gaps: 3
  Percent Similarity: 77.551  Percent Identity: 77.551

```

alignment_block:

US-09-306-689-10 x PCT-US00-22121-35 ..

Align seg 1/1 to: PCT-US00-22121-35 from: 1 to: 40

```

1 CAGCATGTGAGCTACGGCCCTGGCCCTGGCAGCGGTTCTCAAGATTGAG 50
|||||
1 GlnHSTPserTyrGlyLeuArpProGly.....GlnHSTPse 14
51 CTACGGCCTGCGTCCGGGTGGCTCTAGCCAGCATTTGGAGCTACGGCTTC 100
|||||
14 rTyGlyLeuArpProGly.....GlnHSTPserTyrGlyLeuA 28
101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCTGCGCGGT 147
|||||
28 rGProGly.....GlnHSTPserTyrGlyLeuArpProGly 40

```

seq_name: /cgn2_6/plodata/2/paa/US090_COMB.pep:US-09-026-276-35

seq_documentation_block:

: Sequence 35, Application US/09026276

: GENERAL INFORMATION:

: APPLICANT: Kenten, John H

: APPLICANT: Tramoncano, Alfonso

: APPLICANT: Pilon, Aprile L

: APPLICANT: Lohnas, Gerald L

: APPLICANT: Roberts, Steven F

: TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM

: FILE REFERENCE: U.S. Patent Application No. 09/026,276

: CURRENT APPLICATION NUMBER: US/09/026,276

: NUMBER OF SEQ ID NOS: 35

: SOFTWARE: PatentIn Ver. 2.0

: SEQ ID NO 35

: LENGTH: 40

: TYPE: PRT

: ORGANISM: Porcine

: US-09-026-276-35

```

alignment_scores:
  Quality: 199.50      Length: 49
  Ratio: 5.250        Gaps: 3
  Percent Similarity: 77.551  Percent Identity: 77.551

```

alignment_block:

US-09-306-689-10 x US-09-026-276-35 ..

Align seg 1/1 to: US-09-026-276-35 from: 1 to: 40

```

1 CAGCATGTGAGCTACGGCCCTGGCCCTGGCAGCGGTTCTCAAGATTGAG 50
|||||
1 GlnHSTPserTyrGlyLeuArpProGly.....GlnHSTPse 14
51 CTACGGCCTGCGTCCGGGTGGCTCTAGCCAGCATTTGGAGCTACGGCTTC 100
|||||
14 rTyGlyLeuArpProGly.....GlnHSTPserTyrGlyLeuA 28

```

```

101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCTGCTCCGGGT 147
|||||
28 rprgoly.....GlnHstprserTyrglyLeuArgrProgly 40

```

seq_name: /cgn2.6/ptodata/2/paa/US093_COMB.pep:US-09-374-721A-35

```

seq_documentation_block:
; Sequence 35, Application US/09374721A
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H.
; APPLICANT: Roberts, Steven
; APPLICANT: Lohas, Gerald
; TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: CIP OF IGN-9601
; CURRENT APPLICATION NUMBER: US/09/374,721A
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: 09/026,276
; PRIOR FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 35
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: polypeptide
; OTHER INFORMATION: antigen
US-09-374-721A-35

```

```

alignment_scores:
Quality: 199.50      Length: 49
Ratio: 5.250         Gaps: 3
Percent Similarity: 77.551 Percent Identity: 77.551

```

alignment_block:

US-09-306-689-10 x US-09-374-721A-35 ..

Align seg 1/1 to: US-09-374-721A-35 from: 1 to: 40

```

1 CAGCATTTGAGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
1 GlnHstprserTyrglyLeuArgrProgly.....GlnHstprse 14
14 rTyrglyLeuArgrProgly.....GlnHstprserTyrglyLeuArgrProgly 28
51 CTACGGCTGCGTCCGGGTGAGCTCTACGCCAGATTGAGCTAGAGCTCCGGT 100
|||||
14 rTyrglyLeuArgrProgly.....GlnHstprserTyrglyLeuArgrProgly 28
101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCTGCTCCGGGT 147
|||||
28 rprgoly.....GlnHstprserTyrglyLeuArgrProgly 40

```

seq_name: /cgn2.6/ptodata/2/paa/PCTUS_COMB.pep:PCT-US00-22121-34

```

seq_documentation_block:
; Sequence 34, Application PC/TUS0022121
; GENERAL INFORMATION:
; APPLICANT: Proteinix Company
; TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: IGN-2004WO
; CURRENT APPLICATION NUMBER: PCT/US00/22121
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 09/026,276
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 09/374,721
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 34
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: polypeptide
; OTHER INFORMATION: antigen
PCT-US00-22121-34

```

```

alignment_scores:
Quality: 199.50      Length: 49
Ratio: 5.250         Gaps: 3
Percent Similarity: 77.551 Percent Identity: 77.551

```

alignment_block:

US-09-306-689-10 x PCT-US00-22121-34 ..

Align seg 1/1 to: PCT-US00-22121-34 from: 1 to: 41

```

1 CAGCATTTGAGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
1 GlnHstprserTyrglyLeuArgrProgly.....GlnHstprse 14
51 CTACGGCTGCGTCCGGGTGAGCTCTACGCCAGATTGAGCTAGAGCTCCGGT 100
|||||
14 rTyrglyLeuArgrProgly.....GlnHstprserTyrglyLeuArgrProgly 28
101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCTGCTCCGGGT 147
|||||
28 rprgoly.....GlnHstprserTyrglyLeuArgrProgly 40

```

seq_name: /cgn2.6/ptodata/2/paa/US090_COMB.pep:US-09-026-276-34

```

seq_documentation_block:
; Sequence 34, Application US/09026276
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H.
; APPLICANT: Tramoniano, Alfonso
; APPLICANT: Pilon, April L.
; APPLICANT: Lohas, Gerald L.
; APPLICANT: Roberts, Steven F.
; TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: U.S. Patent Application No. 09/026,276
; CURRENT APPLICATION NUMBER: US/09/026,276
; PRIOR FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 34
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Porcine
US-09-026-276-34

```

```

alignment_scores:
Quality: 199.50      Length: 49
Ratio: 5.250         Gaps: 3
Percent Similarity: 77.551 Percent Identity: 77.551

```

alignment_block:

US-09-306-689-10 x US-09-026-276-34 ..

Align seg 1/1 to: US-09-026-276-34 from: 1 to: 41

```

1 CAGCATTTGAGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
1 GlnHstprserTyrglyLeuArgrProgly.....GlnHstprse 14
51 CTACGGCTGCGTCCGGGTGAGCTCTACGCCAGATTGAGCTAGAGCTCCGGT 100
|||||
14 rTyrglyLeuArgrProgly.....GlnHstprserTyrglyLeuArgrProgly 28
101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGGCTGCTCCGGGT 147
|||||
28 rprgoly.....GlnHstprserTyrglyLeuArgrProgly 40

```


XX 24-JAN-1996; 96WO-CA00049.
 PF
 XX
 PR 10-FEB-1995; 95US-0387156.
 XX
 PA (UYSA-) UNIV SASKATCHEWAN.
 PI
 PI Manns JG, Potter AA;
 XX
 DR MPI: 1996-184447/38.
 DR N-PSDB; T37178.
 XX
 PT Gonadotropin-releasing hormone multimer fusion proteins - with
 PT leukotoxin polypeptide for increased immunogenicity, useful in
 PT antiferility vaccine prodn.
 PS
 PS Example 2; Fig 1B; 87pp; English.
 XX
 CC A synthetic DNA sequence (T37178) codes for a gonadotropin
 CC releasing hormone (GNRH) tetramer (W03944), in which the 4 GNRH
 CC repeat units are separated by spacers designed to increase
 CC immunogenicity. The DNA sequence was incorporated into vector
 CC PAA332 (ATCC 68283), which contains a truncated leukotoxin
 CC gene (LKT 352) derived from Pasteurella haemolytica, to
 CC give plasmid PCB113 (T37176). Escherichia coli transformants
 CC produce an LKT-GNRH fusion protein (see also W03942 and W03943)
 CC useful as a vaccine for fertility control of domestic or farm animals.
 CC
 XX Sequence 49 AA;
 SQ

Query Match 100.0%; Score 290; DB 17; Length 49;
 Best Local Similarity 100.0%; Pred. No. 1.7e-25;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSGSQDWSYGLRPG 49
 DB 1 qhwsyglrpgsgsqdwsyglrpgssqhwsyglrpgsgsqdwsyglrpg 49

RESULT 2
 W79567
 ID W79567 standard; Protein; 49 AA.
 XX
 AC W79567;
 XX
 DT 24-DEC-1998 (first entry)
 XX
 DE GNRH-2.
 XX
 KW Gonadotropin releasing hormone; GNRH; decapeptide; hypothalamus; spacer;
 KW LH; luteinising hormone; FSH; follicle stimulating hormone; vertebrate;
 KW pyroglu; chimera; leukotoxin polypeptide; multimer; vaccine; tumour;
 KW immunogenic.
 XX
 OS Synthetic.
 XX
 PN WO9806848-A1.
 XX
 PD 19-FEB-1998.
 XX
 PF 08-AUG-1997; 97WO-CA00559.
 XX
 PR 09-AUG-1996; 96US-0694865.
 XX
 PA (UYSA-) UNIV SASKATCHEWAN.
 PI
 PI Manns JG, Potter AA;
 XX
 DR MPI: 1998-159540/14.
 DR N-PSDB; V61529.
 XX
 PT Chimeric protein of leukotoxin and gonadotropin releasing hormone

PT useful for, e.g. preparation of vaccines for reduction of incidence
 PT of mammary tumours in mammals
 XX
 PS Disclosure; Figure 1B; 118pp; English.
 XX
 CC The present sequence represents a recombinantly produced or chemically
 CC synthesised gonadotropin releasing hormone-2 (GNRH-2) polypeptide, which
 CC contains four copies of the GNRH decapeptide (V61528) and triplet amino
 CC acid spacers between each of these sequences. This decapeptide is
 CC secreted naturally by the hypothalamus which controls release of both the
 CC luteinising hormone (LH) and the follicle stimulating hormone (FSH) in
 CC vertebrates. This sequence, as compared to the native peptide, has been
 CC found to have an N-terminal Gln rather than a pyroglu residue, and also
 CC contains substitutions at amino acid residues 15 and 41, whereby His is
 CC replaced by Asp. This produces an alternating multimeric GNRH sequence
 CC which is highly immunogenic that can be used in the construction of a
 CC chimeric protein that comprises a leukotoxin polypeptide, several
 CC multimers, and the GNRH sequence. The chimeric protein can be used as a
 CC vaccine to help reduce the incidence of mammary tumours in a mammalian
 CC individual.
 CC
 XX Sequence 49 AA;
 SQ

Query Match 100.0%; Score 290; DB 19; Length 49;
 Best Local Similarity 100.0%; Pred. No. 1.7e-25;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSGSQDWSYGLRPG 49
 DB 1 qhwsyglrpgsgsqdwsyglrpgssqhwsyglrpgsgsqdwsyglrpg 49

RESULT 3
 W61542
 ID W61542 standard; Protein; 49 AA.
 XX
 AC W61542;
 XX
 DT 27-OCT-1998 (first entry)
 XX
 DE Peptide hormone GNRH-2 decapeptide (4 copies) fragment.
 XX
 KW GNRH; gonadotropin releasing hormone peptide hormone; leukotoxin;
 KW immunisation; endogenous molecule; vaccine; ear; immunogen; carrier;
 KW immune response; hormone receptor; cancerous cell; domestic animal;
 KW porcine; bovine; luteinizing hormone; follicle stimulating hormone;
 KW immunocastate.
 XX
 OS Synthetic.
 XX
 PN WO9834639-A1.
 XX
 PD 13-AUG-1998.
 XX
 PF 04-FEB-1998; 98WO-CA00059.
 XX
 PR 05-FEB-1997; 97US-0036883.
 XX
 PA (BIOS-) BIOSSTAR INC.
 PI
 PI Acres SD, Harland R, Manns JG;
 XX
 DR MPI: 1998-446952/38.
 DR N-PSDB; V45190.
 XX
 PT Immunisation against endogenous molecules by administering vaccine
 PT to ear - useful to elicit efficient and uniform immune response
 PT against e.g. gonadotropin releasing hormone to immunocastate pigs
 PT and cattle
 XX
 XX Example 1; Fig 1B; 61pp; English.
 XX

CC This represents the amino acid sequence of the gonadotropin releasing
 CC hormone (GnRH-2) decapeptide fragment used in the chimeric leukotoxin-
 CC GnRH polypeptide gene fusions. This is used to exemplify the method of
 CC invention of immunisation against endogenous molecules by administering
 CC a vaccine which comprises an immunogen and a carrier to the ear of the
 CC mammal. The method is useful for eliciting an efficient and uniform
 CC immune response to block or suppress the activity of an endogenous
 CC hormone, hormone receptor, agonist or antagonist in a vaccinated subject,
 CC or to elicit an immune response against a targeted endogenous cell type
 CC (e.g. a cancerous or otherwise diseased cell). It is especially useful
 CC to reduce the levels of GnRH in domestic animals, especially in porcine
 CC or bovine species. The use of GnRH immunogens in the vaccine reduces the
 CC levels of luteinizing hormone and follicle stimulating hormone and helps
 CC in immunocastrating the animal. Administration of vaccine compositions to
 CC the ear instead of intramuscular administration into the neck increases
 CC the efficiency of vaccination of mammals against endogenous immunogens,
 CC and may increase uniformity of vaccine presentation since the ear is
 CC relatively uniform from animal to animal.

XX Sequence 49 AA:

Query Match 100.0%; Score 290; DB 19; Length 49;
 Best Local Similarity 100.0%; Pred. No. 1.7e-25;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OHWSYGLRPGSGSDMSYGLRPGSSQHMWSYGLRPGSGSDMSYGLRPG 49
 DB 1 qhwsyglrpgsgsdmsyglrpgsgsdmsyglrpgsgsdmsyglrpg 49

RESULT 4
 Y58363
 ID Y58363 standard; Protein; 49 AA.

XX Y58363;

DT 27-MAR-2000 (first entry)

XX Four-copy gonadotropin-releasing hormone (GnRH) multimer.

KW GnRH multimer; gonadotropin-releasing hormone; immunosterilisation;

KW immunocastration; vaccine; feline; canine; equine; cervine; ds

XX Mammalia.

OS Synthetic.

XX MO9962545-A2.

PN 09-DEC-1999.

PD 09-DEC-1999.

PF 28-MAY-1999; 99WO-CA00493.

XX 04-JUN-1998; 98US-0088024.

PR 06-MAY-1999; 99US-0306689.

XX (BIOS-) BIOSTAR INC.

PA (BIOS-) BIOSTAR INC.

XX Robbins SC;

PI Robbins SC;

XX WPI; 2000-086857/07.

DR N-PSDB; 255702.

XX Hormone immunogens, analogues or antibodies used to manufacture

PT vaccines for suppression of reproductive behavior and fertility in

PT vertebrates -

XX Claim 5; Fig 5B; 88pp; English.

PS Claim 5; Fig 5B; 88pp; English.

XX This sequence represents a four-copy gonadotropin-releasing

CC hormone (GnRH) multimer, where the second and fourth GnRH sequence

CC have a His to Asp substitution at position 2 of the GnRH sequence.

CC The invention relates to GnRH immunogens, analogues or antibodies

CC that cross-react with endogenous GnRH of a vertebrate. A specifically
 CC claimed immunogenic fusion protein (Y58361) comprises, in the N to
 CC C-terminal direction, a synthetic peptide sequence (Y58364), an eight
 CC copy GnRH multimer (composed of two copies of the 4xGnRH multimer
 CC sequence of Y58363), the LKT protein (which functions as a carrier
 CC protein), and a second eight copy GnRH multimer. The fusion protein may
 CC be used in a vaccine composition for prepubertal administration to a
 CC vertebrate subject to result in prolonged suppression of reproductive
 CC behaviour and/or fertility. GnRH immunogens, analogues or antibodies are
 CC used to manufacture a composition or vaccine for immunosterilisation or
 CC immunocastration of feline, canine, equine or cervine subjects.
 CC The vaccines are used to suppress reproductive behaviour and/or
 CC fertility for at least 10 months. The prepubertal administration
 CC results in a prolonged, long-term suppression of testicular development
 CC and/or function in males, or a prolonged, long-term suppression of
 CC ovarian development and/or function in females. The methods provide a
 CC viable and desirable alternative to surgical forms of sterilisation that
 CC are currently used.

XX Sequence 49 AA:

Query Match 100.0%; Score 290; DB 21; Length 49;
 Best Local Similarity 100.0%; Pred. No. 1.7e-25;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OHWSYGLRPGSGSDMSYGLRPGSSQHMWSYGLRPGSGSDMSYGLRPG 49
 DB 1 qhwsyglrpgsgsdmsyglrpgsgsdmsyglrpgsgsdmsyglrpg 49

RESULT 5
 Y58135
 ID Y58135 standard; Protein; 49 AA.

XX Y58135;

DT 07-MAR-2000 (first entry)

XX GnRH analogue multimer, containing four copies of the GnRH analogue.

KW Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;

KW antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;

KW non-androgenic; steroid; reduction; weight gain; muscle distribution;

KW fat distribution; male pattern; boar taint; flavour; impalement;

XX reliable; immunocastration; meat production.

OS Synthetic.

XX Mammalia.

PN WO9956771-A2.

PD 11-NOV-1999.

PF 05-MAY-1999; 99WO-CA00360.

XX 05-MAY-1998; 98US-0084217.

XX (BIOS-) BIOSTAR INC.

PA (BIOS-) BIOSTAR INC.

XX Manns JG, Acres SD, Harland R;

PI Manns JG, Acres SD, Harland R;

XX WPI; 2000-062125/05.

DR N-PSDB; 246402.

XX Production of uncastrated male food animals using vaccines -

PT Example 1; Fig 2B; 87pp; English.

PS Example 1; Fig 2B; 87pp; English.

XX This sequence represents four copies of a gonadotropin

CC releasing hormone (GnRH) analogue, DNA encoding which was

CC used in the construction of a chimeric GnRH-leukotoxin (LKT)

CC fusion gene (246400). This fusion gene encodes a GnRH-LKT fusion

CC protein which may be used as a vaccine. The LKT portion of the protein
 CC acts to enhance the immunogenicity of the GnRH portion. The invention
 CC relates to a method of using two GnRH immunogen vaccines to produce
 CC unestrated male animals for meat production, one vaccination prior to
 CC or during the fattening period to reduce circulating testosterone levels,
 CC and the second vaccination about 2-8 weeks before slaughter to
 CC substantially reduce androgenic and/or non-androgenic steroids. The
 CC invention is used to produce food animals that exhibit the weight gain
 CC and muscle/fat distribution of male animals without the problems
 CC associated with male animals. Such problems include "boar taint", a
 CC urine-like odour found in cooked meat of unestrated pigs which is
 CC caused by steroids stored in the tissues, and similar flavour
 CC impairments in the meat of other intact male animals. The invention is
 CC more reliable than prior art immunocastration techniques.

SO Sequence 49 AA;

Query Match 100.0%; Score 290; DB 21; Length 49;
 Best Local Similarity 100.0%; Pred. No. 1.7e-25;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHMSYGLRPGSGSDMSYGLRPG 49
 DB 1 qhwsyglrpgsgsdmsyglrpgssqhmsyglrpgsgsdmsyglrpg 49

RESULT 6

ID W03943 standard; Protein; 544 AA.

AC W03943;

DT 20-NOV-1996 (first entry)

DE LKT-GnRH protein fusion from PCB111.

XX Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;

KW fusion protein; immunogen; vaccine; fertility control;

XX KW contraceptive; sterilisation.

OS Chimeric Pasteurella haemolytica A1 strain B122;

XX OS Chimeric synthetic.

FT Key Location/Qualifiers

FT Domain 1..493 /label= LKT

FT Domain 494..544 /label= GnRH_repeat_domain

PN W09624675-A1.

PD 15-AUG-1996.

PF 24-JAN-1996; 96WO-CA00049.

PR 10-FEB-1995; 95US-0387156.

XX (UYSA-) UNIV SASKATCHEWAN.

PI Manns JG, Potter AA;

DR WPI: 1996-384447/38.

DR N-PSDB: T37177.

XX Gonadotropin-releasing hormone multimer fusion proteins - with

PT leukotoxin polypeptide for increased immunogenicity, useful in

PT antiferility vaccine prodn.

XX Claim 8; Fig 7A-7E; 87pp; English.

XX A chimeric protein (W03943) is composed of a fusion between

CC a truncated leukotoxin (LKT-111) from Pasteurella haemolytica

CC and a 4-copy gonadoliberin-releasing hormone (GnRH) repeat
 CC sequence (see also W03944). It is the product of a chimeric
 CC gene (T37177) produced by deleting an approx. 1300 bp sequence
 CC from PCB113 (see also T37176) coding for amino acids 352-784
 CC of LKT-352. Recombinant plasmid PCB111 (LKT 111:4 copy GnRH,
 CC ATCC 69748) was obtd. Escherichia coli transformants produced
 CC the chimeric protein, which is useful as a vaccine for fertility
 CC control, esp. immunological sterilisation of domestic or farm
 CC animals.

SO Sequence 544 AA;

Query Match 100.0%; Score 290; DB 17; Length 544;
 Best Local Similarity 100.0%; Pred. No. 2.3e-24;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHMSYGLRPGSGSDMSYGLRPG 49
 DB 494 qhwsyglrpgsgsdmsyglrpgssqhmsyglrpgsgsdmsyglrpg 542

RESULT 7

ID W79570 standard; Protein; 544 AA.

AC W79570;

DT 24-DEC-1998 (first entry)

DE LKT-GnRH chimeric protein.

XX Chimeric; PCB111; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;

KW cytotoxic activity; antigen presentation; immune response; vaccine;

XX KW tumour.

OS Synthetic.

PN W09806848-A1.

PD 19-FEB-1998.

PF 08-AUG-1997; 97WO-CA00559.

PR 09-AUG-1996; 96US-0694865.

XX (UYSA-) UNIV SASKATCHEWAN.

PI Manns JG, Potter AA;

DR WPI: 1998-159540/14.

DR N-PSDB: V61532.

XX Chimeric protein of leukotoxin and gonadotropin releasing hormone

PT useful for, e.g. preparation of vaccines for reduction of incidence

PT of mammary tumours in mammals

XX Disclosure; Figure 7.1-5; 118pp; English.

XX The present sequence represents the LKT-GnRH chimeric protein from

CC PCB111. This plasmid contains the LKT 111 polypeptide fused to

CC four copies of the GnRH peptide. This chimera lacks cytotoxic activity

CC which enables there to be an increase in antigen presentation and thus an

CC optimal immune response. The removal of this region also enables the

CC truncated LKT to be expressed at much higher levels and allows the amount

CC of antigen administered to be reduced. This chimeric protein comprises a

CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The

CC chimeric protein can be used as a vaccine to help reduce the incidence of

CC mammary tumours in a mammalian individual.

SO Sequence 544 AA;

Query Match 100.0%; Score 290; DB 19; Length 544;
 Best Local Similarity 100.0%; Pred. No. 2.3e-24;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHSYGLRPGSGSDMSYGLRPG 49
 |||||||
 DB 494 qhwsyglrpgsgsqdwsyglrpgsgsqhwsyglrpgsgsqdwsyglrpg 542

RESULT 8

W79573
 ID W79573 standard; Protein; 695 AA.

AC W79573;

DT 24-DEC-1998 (first entry)

DE LKT-GnRH chimeric protein.

XX Chimera: PCB122; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;

KW cytotoxic activity; antigen presentation; Immune response; vaccine;

XX tumour.

OS Synthetic.

PN MO9806848-A1.

PD 19-FEB-1998.

PF 08-AUG-1997; 97WO-CA00559.

PR 09-AUG-1996; 96US-0694865.

XX (UYSA-) UNIV SASKATCHEWAN.

PI Manns JG, Potter AA;

DR WPI: 1998-159540/14.

DR N-PSDB: V61535

PT Chimeric protein of leukotoxin and gonadotropin releasing hormone

PT useful for, e.g. preparation of vaccines for reduction of incidence

PT of mammary tumours in mammals

XX Claim 9; Figure 9.1-6; 11pp; English.

XX The present sequence represents the LKT-GnRH chimeric protein from

CC PCB122. This plasmid contains the LKT 111 polypeptide fused to sixteen

CC copies of the GnRH peptide. In the pattern of: 8 copies of GnRH-LKT 111-8

CC copies of GnRH. This chimera lacks cytotoxic activity which enables

CC there to be an increase in antigen presentation and thus an optimal

CC immune response. The removal of this region also enables the truncated

CC LKT to be expressed at much higher levels and allows the amount of

CC antigen administered to be reduced. This chimeric protein comprises a

CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The

CC chimeric protein can be used as a vaccine to help reduce the incidence of

XX mammary tumours in a mammalian individual.

SO Sequence 695 AA;

Query Match 100.0%; Score 290; DB 19; Length 695;

Best Local Similarity 100.0%; Pred. No. 2.9e-24;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHSYGLRPGSGSDMSYGLRPG 49
 |||||||
 DB 9 qhwsyglrpgsgsqdwsyglrpgsgsqhwsyglrpgsgsqdwsyglrpg 57

RESULT 9

Y58361
 ID Y58361 standard; Protein; 695 AA.

XX Y58361;

AC 27-MAR-2000 (first entry)

DT Leukotoxin/gonadotropin-releasing hormone fusion protein.

DE Leukotoxin; gonadotropin-releasing hormone; GnRH; immunostimulation;

XX immunosuppression; vaccine; feline; canine; equine; cervine.

OS Chimeric - Pasteurella haemolytica.

OS Chimeric - Mammalia.

PN MO9962545-A2.

PD 28-MAY-1999; 99WO-CA00493.

PF 04-JUN-1998; 98US-0088024.

PR 06-MAY-1999; 99US-0306689.

XX (BIO-) BIOSTAR INC.

PI Robbins SC;

DR WPI: 2000-086857/07.

DR N-PSDB: Z55700.

PT Hormone immunogens, analogues or antibodies used to manufacture

PT vaccines for suppression of reproductive behavior and fertility in

PT vertebrates

XX Claim 20; Fig 6A-6F; 88pp; English.

XX This sequence represents a fusion protein comprising gonadotropin-

CC releasing hormone (GnRH) immunogens and a Pasteurella

CC haemolytica leukotoxin (LKT) protein. The fusion protein comprises, in

CC the N to C-terminal direction, a synthetic peptide sequence (Y58364), an

CC eight copy GnRH multimer (composed of two copies of the 4xGnRH multimer

CC sequence of Y58363), the LKT protein (which functions as a carrier

CC protein), and a second eight copy GnRH multimer. The fusion protein may

CC be used in a vaccine composition for prepubertal administration to a

CC vertebrate subject to result in prolonged suppression of reproductive

CC behaviour and/or fertility. GnRH immunogens, analogues or antibodies

CC that cross-react with endogenous GnRH of a vertebrate subject are used

CC to manufacture a composition or vaccine for immunosuppression or

CC immunosuppression of feline, canine, equine or cervine subjects.

CC The vaccines are used to suppress reproductive behaviour and/or

CC fertility for at least 10 months. The prepubertal administration

CC results in a prolonged, long-term suppression of testicular development

CC and/or function in males, or a prolonged, long-term suppression of

CC ovarian development and/or function in females. The methods provide a

CC viable and desirable alternative to surgical forms of sterilisation that

XX are currently used.

SO Sequence 695 AA;

Query Match 100.0%; Score 290; DB 21; Length 695;

Best Local Similarity 100.0%; Pred. No. 2.9e-24;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHSYGLRPGSGSDMSYGLRPG 49
 |||||||
 DB 9 qhwsyglrpgsgsqdwsyglrpgsgsqhwsyglrpgsgsqdwsyglrpg 57

RESULT 10

Y58133
 ID Y58133 standard; Protein; 695 AA.

AC Y58133;

```

XX 07-MAR-2000 (first entry)
XX Gonadotropin releasing hormone-leukotoxin fusion protein.
DE
XX
XX Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;
XX antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;
XX non-androgenic; steroid; reduction; weight gain; muscle distribution;
XX fat distribution; male pattern; boar taint; flavour; impairment;
XX reliable; immunocastration; meat production.
XX
XX Chimeric - Mammalia.
OS Chimeric - Pasteurella haemolytica.
XX
XX WO956771-A2.
XX
XX 11-NOV-1999.
XX
XX 05-MAY-1999; 99WO-CA00360.
XX
XX 05-MAY-1998; 98US-0084217.
XX
XX (BIOS-) BIOSSTAR INC.
XX
XX Manns JG, Acres SD, Harland R;
XX
XX WPI; 2000-062125/05.
XX
XX N-PSDB; 246400.
XX
XX Production of uncastrated male food animals using vaccines -
XX
XX Claim 22; Fig 3A-3F; 87pp; English.
XX
XX This sequence represents a chimeric gonadotropin releasing
XX hormone (GnRH)-leukotoxin (LKT) fusion protein, which may be
XX used as a vaccine. The LKT portion of the protein acts to enhance
XX the immunogenicity of the multimeric GnRH portion (Y58135). The invention
XX relates to a method of using two GnRH immunogen vaccines to produce
XX uncastrated male animals for meat production, one vaccination prior to
XX or during the fattening period to reduce circulating testosterone levels,
XX and the second vaccination about 2-8 weeks before slaughter to
XX substantially reduce androgenic and/or non-androgenic steroids. The
XX invention is used to produce food animals that exhibit the weight gain
XX and muscle/fat distribution of male animals without the problems
XX associated with male animals. Such problems include "boar taint", a
XX urine-like odour found in cooked meat of uncastrated pigs which is
XX caused by steroids stored in the tissues, and similar flavour
XX impairments in the meat of other intact male animals. The invention is
XX more reliable than prior art immunocastration techniques.
XX
XX Sequence 695 AA:
SQ

```

Query Match 100.0%; Score 290; DB 21; Length 695;
 Best Local Similarity 100.0%; Pred. No. 2.9e-24;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OHMSYGLRPGSGSDMSYGLRPGSSOHMSYGLRPGSGSDMSYGLRPG 49
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 9 qhwsyglirpssgdsyglirpssgshwsyglirpssgdsyglirp 57

RESULT 11
 ID W03942
 W03942 standard; Protein: 977 AA.
 XX
 XX W03942;
 XX
 XX 20-NOV-1996 (first entry)
 XX
 XX LKT-GnRH protein fusion from PCB113.
 DE
 XX Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
 XX

```

KW fusion protein; immunogen; vaccine; fertility control;  

KW contraceptive; sterilisation.  

XX  

XX Chimeric Pasteurella haemolytica A1 strain B122;  

OS Chimeric synthetic.  

XX  

XX Key Location/Qualifiers  

XX Domain 1..929  

XX Domain /label= LKT  

XX Domain 927..977  

XX Domain /label= GnRH_repeat_domain  

XX  

XX WO9624675-A1.  

XX  

XX 15-AUG-1996.  

XX  

XX 24-JAN-1996; 96WO-CA00049.  

XX  

XX 10-FEB-1995; 95US-0387156.  

XX  

XX (UYSA-) UNIV SASKATCHEWAN.  

XX  

XX Manns JG, Potter AA;  

XX  

XX WPI; 1996-384447/38.  

XX  

XX N-PSDB; T37176.  

XX  

XX Gonadotropin-releasing hormone multimer fusion proteins - with  

XX leukotoxin polypeptide for increased immunogenicity, useful in  

XX antifertility vaccine prodn.  

XX  

XX Claim 7; Fig 5A-5H; 87pp; English.  

XX  

XX A chimeric protein (W03942) is composed of a fusion between  

XX a truncated leukotoxin (LKT-352) from Pasteurella haemolytica (see  

XX also W03945) and a 4-copy gonadotropin-releasing hormone (GnRH)  

XX repeat sequence (see also W03944). It is the product of a  

XX chimeric gene (T37176) produced by ligating a synthetic sequence  

XX for the 4-copy GnRH into vector pAA352 (ATCC 68283), which carries  

XX the LKT-352 gene. Recombinant plasmid PCB113 (LKT 352:4 copy  

XX GnRH; ATCC 69749) was obt'd. Escherichia coli transformants  

XX produced the chimeric protein, which is useful as a vaccine for  

XX fertility control, esp. immunological sterilisation of  

XX domestic or farm animals.  

XX  

XX Sequence 977 AA:  

SQ

```

Query Match 100.0%; Score 290; DB 17; Length 977;
 Best Local Similarity 100.0%; Pred. No. 4.2e-24;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OHMSYGLRPGSGSDMSYGLRPGSSOHMSYGLRPGSGSDMSYGLRPG 49
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 927 qhwsyglirpssgdsyglirpssgshwsyglirpssgdsyglirp 975

RESULT 12
 ID W79569
 W79569 standard; Protein: 977 AA.
 XX
 XX W79569;
 XX
 XX 24-DEC-1998 (first entry)
 XX
 XX LKT-GnRH chimeric protein.
 DE
 XX Chimeric; PCB113; LKT 352; GnRH; Gonadotropin releasing hormone; multimer;
 XX cytotoxic activity; antigen presentation; immune response; vaccine;
 XX
 XX Synthetic.
 XX

PM WO9806848-A1.
 XX 19-FEB-1998.
 PD 08-AUG-1997; 97WO-CA00559.
 XX 09-AUG-1996; 96US-0694865.
 PR (UYSA-) UNIV SASKATCHEWAN.
 XX Manns JG, Potter AA;
 PI WPI: 1998-159540/14.
 DR N-PSDB; V61531.
 XX Chimeric protein of leukotoxin and gonadotropin releasing hormone
 PT useful for, e.g. preparation of vaccines for reduction of incidence
 PT of mammary tumours in mammals
 PS Disclosure: Figure 5.1-8; 118pp; English.
 XX The present sequence represents the LKT-GnRH chimeric protein from
 CC pCbl13. This plasmid contains the LKT 352 polypeptide (W79568) fused to
 CC four copies of the GnRH peptide. This chimera lacks cytotoxic activity
 CC which enables there to be an increase in antigen presentation and thus an
 CC optimal immune response. The removal of this region also enables the
 CC truncated LKT to be expressed at much higher levels and allows the amount
 CC of antigen administered to be reduced. This chimeric protein comprises a
 CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The
 CC chimeric protein can be used as a vaccine to help reduce the incidence of
 CC mammary tumours in a mammalian individual.
 XX Sequence 977 AA;
 SQ
 Query Match 100.0%; Score 290; DB 19; Length 977;
 Best Local Similarity 100.0%; Pred. No. 4.2e-24;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 OHWSYGLRPGSGQDWSYGLRPGSSQHWSYGLRPGSGQDWSYGLRPG 49
 Db 927 qhwsyglrpgsgqdwmsyglrpgsgsqhwsyglrpgsgsqdwmsyglrpg 975
 RESULT 13
 Y31183
 ID Y31183 standard; peptide; 40 AA.
 XX Y31183;
 AC 28-OCT-1999 (first entry)
 DT Ubiquitin fusion protein GnRH fragment 2.
 DE Ubiquitin fusion protein GnRH fragment 2.
 XX Ubiquitin; immunocastration; fusion protein; heat shock protein; epitope;
 KW immune response stimulation; vaccine; T cell; viral; infection; cancer;
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;
 KW pulmonary infection; respiratory infection; scaffold; anti-self; pig;
 KW steriodogenesis; gamete maturation; prostate; breast; castration; TNF;
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;
 KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;
 KW fertility; sperm protein; growth rate; antibody; detection; GnRH.
 XX Unidentified.
 OS WO9942472-A1.
 XX 26-AUG-1999.
 PD 26-JAN-1999; 99WO-US01588.
 PF 19-FEB-1998; 98US-0026276.
 XX

PA (IGEN-) IGEN INT INC.
 XX Kenten JH, Lohmas GL, Pilon AL, Roberts SF, Tramontano A;
 PI WPI: 1999-518582/43.
 DR Epitope-containing fusion proteins used to generate a highly
 XX specific immune responses
 PT Claim 83; Page 43; 67pp; English.
 PS This invention describes a novel fusion protein, comprising a heat shock
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner
 CC which is useful for the stimulation of a highly specific immune response
 CC when administered to an animal. The protein of the invention may be
 CC post-translationally modified (e.g. by the addition of fatty acids to
 CC enhance immunogenicity). The fusion proteins of the invention can be
 CC used as vaccines to induce an immune response. When a T cell epitope is
 CC attached, they can be used for control of viral infections, bacterial
 CC infections, parasitic infection and cancer. The fusion proteins can be
 CC used in pharmaceutical compositions for the treatment of gastrointestinal
 CC diseases, pulmonary infections, respiratory infections, and HIV
 CC infections. The use of ubiquitin as a scaffold is also useful for the
 CC generation of anti-gonadotropin releasing hormone antibodies which result
 CC in the suppression of luteinizing hormone and follicle stimulating
 CC hormone. This indirectly suppresses steriodogenesis and gamete maturation
 CC in males and females. This type of anti-self response in humans is useful
 CC in the treatment of prostate cancer and breast cancer. In livestock, the
 CC ability to stimulate an anti-self response provides a simple alternative
 CC to physical castration. Immunocastration of pigs is a better alternative
 CC to physical castration, as it does not result in any of the detrimental
 CC side effects associated with physical castration. Other examples of
 CC diseases and conditions treated with self proteins fused with ubiquitin
 CC are TNF and its epitopes to modulate septic shock, arthritis,
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig
 CC epsilon heavy chain for the control of allergic reactions; chorionic
 CC gonadotropin for fertility control; and sperm proteins for fertility
 CC control. A further use of the fusion proteins is as part of a vaccine to
 CC enhance growth rate and thereby the final weight of the livestock prior
 CC to shipment to market. In addition, the fusion proteins of the invention
 CC can be used to detect and identify antibodies from experimental samples.
 CC This sequence represents a GnRH fragment used in the construction of
 CC a ubiquitin fusion protein described in the method of the invention.
 XX Sequence 40 AA;
 SQ
 Query Match 68.8%; Score 199.5; DB 20; Length 40;
 Best Local Similarity 77.6%; Pred. No. 1e-15; 2; Indels 9; Gaps 3;
 Matches 38; Conservative 0; Mismatches 2;
 QY 1 OHWSYGLRPGSGQDWSYGLRPGSSQHWSYGLRPGSGQDWSYGLRPG 49
 Db 1 qhwsyglrpg---qhwsyglrpg---qhwsyglrpg---qhwsyglrpg 40
 RESULT 14
 Y31182
 ID Y31182 standard; peptide; 41 AA.
 XX Y31182;
 AC 28-OCT-1999 (first entry)
 DT Ubiquitin fusion protein GnRH fragment.
 DE Ubiquitin fusion protein GnRH fragment.
 XX Ubiquitin; immunocastration; fusion protein; heat shock protein; epitope;
 KW immune response stimulation; vaccine; T cell; viral; infection; cancer;
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;
 KW pulmonary infection; respiratory infection; scaffold; anti-self; pig;
 KW steriodogenesis; gamete maturation; prostate; breast; castration; TNF;
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;
 KW

KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;
 XX fertility; sperm protein; growth rate; antibody; detection; GnRH.
 OS Unidentified.
 XX
 XX WO9942472-A1.
 PN
 XX
 PD 26-AUG-1999.
 XX
 XX 26-JUN-1999; 99WO-US01588.
 PF
 XX 19-FEB-1998; 98US-0026276.
 PR
 XX (IGEN-) IGEN INT INC.
 PA
 XX
 PI Kenten JH, Lohmas GL, Pilon AL, Roberts SF, Tramontano A;
 DR WPI; 1999-518582/43.
 XX
 XX
 PT Epitope-containing fusion proteins used to generate a highly
 XX specific immune responses
 PS Claim 81; Page 43; 67pp; English.
 XX
 CC This invention describes a novel fusion protein, comprising a heat shock
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner
 CC which is useful for the stimulation of a highly specific immune response
 CC when administered to an animal. The protein of the invention may be
 CC post-translationally modified (e.g. by the addition of fatty acids to
 CC enhance immunogenicity). The fusion proteins of the invention can be
 CC used as vaccines to induce an immune response. When a T cell epitope is
 CC attached, they can be used for control of viral infections, bacterial
 CC infections, parasitic infection and cancer. The fusion proteins can be
 CC used in pharmaceutical compositions for the treatment of gastrointestinal
 CC diseases, pulmonary infections, respiratory infections, and HIV
 CC infections. The use of ubiquitin as a scaffold is also useful for the
 CC presentation and stimulation of anti-self immune responses, e.g.
 CC generation of anti-gonadotropin releasing hormone antibodies which result
 CC in the suppression of luteinizing hormone and follicle stimulating
 CC hormone. This indirectly suppresses steroidogenesis and gamete maturation
 CC in males and females. This type of anti-self response in humans is useful
 CC in the treatment of prostate cancer and breast cancer. In livestock, the
 CC ability to stimulate an anti-self response provides a simple alternative
 CC to physical castration. Immunocastration of pigs is a better alternative
 CC to physical castration, as it does not result in any of the detrimental
 CC side effects associated with physical castration. Other examples of
 CC diseases and conditions treated with self proteins fused with ubiquitin
 CC are TNF and its epitopes to modulate septic shock, arthritis,
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig
 CC epsilon heavy chain for the control of allergic reactions; chorionic
 CC gonadotropin for fertility control; and sperm proteins for fertility
 CC control. A further use of the fusion proteins is as part of a vaccine to
 CC enhance growth rate and thereby the final weight of the livestock prior
 CC to shipment to market. In addition, the fusion proteins of the invention
 CC can be used to detect and identify antibodies from experimental samples.
 CC This sequence represents a GnRH fragment used in the construction of
 CC a ubiquitin fusion protein described in the method of the invention.
 XX
 SO Sequence 41 AA;

Query Match 68.8%; Score 199.5; DB 20; Length 41;
 Best Local Similarity 77.6%; Pred. No. 1,1e-15;
 Matches 38; Conservative 0; Mismatches 2; Indels 9; Gaps 3;

OY 1 QHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPG 49
 Db 1 qhwsyglrpg---qhwsyglrpg---qhwsyglrpg---qhwsyglrpg 40

RESULT 15
 ID R1187 standard; Protein; 323 AA.

XX
 AC R1187;
 XX
 DT 22-MAR-1991 (first entry)
 XX
 DE Plasmid pBT8A859-encoded Tratp-multiple LHRH analogue fusion.
 XX
 XX Tratp protein; luteinizing hormone releasing hormone; fusion protein;
 XX immunological castration.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..20
 FT /label= Tratp signal
 FT Peptide 201..280
 FT /label= 8 LHRH analogues in tandem repeat
 XX
 PN WO9102799-A.
 XX
 XX 07-MAR-1991.
 PD
 XX
 PF 24-AUG-1990; 90WO-AU00373.
 XX
 PR 25-AUG-1989; 89AU-0005979.
 XX
 PA (BIOF-) BIOTECHN AUST PTY L.
 XX
 PI Russell-Jones GJ, Stewart AG, Tsoulis CG;
 XX
 DR WPI; 1991-087282/12.
 DR N-PSDB; Q11021.
 XX
 XX Fusion proteins comprising LHRH analogue and Tratp (analogue) -
 PT useful in vaccine for inhibition or control of reproduction in
 PT vertebrates, esp. domestic animals
 XX
 XX Example 1; Fig 2 and 5; 53pp; English.
 PS
 XX
 CC Plasmid pBT8A859 is a Tratp-LHRH analogue fusion in which 8 copies
 CC of an LHRH analogue have been inserted between amino acids 200 and
 CC 201 of Tratp (Ogata R.T. et al., (1982) J.Bacteriol. 151:819-827).
 CC The plasmid was constructed by two successive additions of DNA
 CC coding for a dimer of LHRH analogue into the SmaI site of pBT8A862
 CC (see Q11020) which all ready carries four copies of the LHRH
 CC sequence. After transformation, colonies with 8 LHRH molecules were
 CC identified. Fusion proteins with multiple inserts generated a higher
 CC anti-LHRH response (as measured by the binding of (125)I-LHRH at a
 CC serum dilution of 1:2000 final) than constructs with a single
 CC insert, in outbred mice and dogs. The fusion proteins can be used to
 CC inhibit reproductive functions in vertebrates.
 CC See also Q10995, Q10997-Q11000, Q11014-Q11020.
 XX
 SO Sequence 323 AA;

Query Match 64.7%; Score 187.5; DB 12; Length 323;
 Best Local Similarity 69.4%; Pred. No. 2e-13;
 Matches 34; Conservative 4; Mismatches 2; Indels 9; Gaps 3;

OY 1 QHWSYGLRPGSSQDMSYGLRPGSSQHWSYGLRPGSSQDMSYGLRPG 49
 Db 201 ehwsyglrpg---ehwsyglrpg---ehwsyglrpg---ehwsyglrpg 240

Search completed: March 2, 2001, 10:53:47
 Job time: 912 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:59:03 ; Search time 201.99 Seconds

(without alignments)
37.240 Million cell updates/sec

Title: US-09-306-689-11

Perfect score: 290
Sequence: 1 QHWSYGLRPGSGSQDMSYGLRG.....SYGLRPGSGSQDMSYGLRPG 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 957798 seqs, 153513742 residues

Total number of hits satisfying chosen parameters: 957798

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA:*
1: /cgn2_6/ptodata/2/paa/PCNUS.COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US05.COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07.COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081.COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082.COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083.COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084.COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085.COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087.COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088.COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089.COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090.COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091.COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092.COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093.COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094.COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095.COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096.COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097.COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098.COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099.COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US100.COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US101.COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US102.COMB.pep:*
27: /cgn2_6/ptodata/2/paa/US103.COMB.pep:*
28: /cgn2_6/ptodata/2/paa/US104.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290	100.0	49	US-09-019-010-4	Sequence 4, Appl1
2	290	100.0	49	US-09-249-447A-7	Sequence 7, Appl1
3	290	100.0	49	US-09-305-924-11	Sequence 11, Appl1
4	290	100.0	49	US-09-306-689-11	Sequence 11, Appl1
5	290	100.0	49	US-09-383-912-4	Sequence 4, Appl1

6	290	100.0	544	17	US-09-383-912-10	Sequence 10, Appl1
7	290	100.0	695	17	US-09-305-924-13	Sequence 13, Appl1
8	290	100.0	695	17	US-09-306-689-13	Sequence 13, Appl1
9	290	100.0	699	17	US-09-383-912-16	Sequence 16, Appl1
10	290	100.0	977	17	US-09-383-912-8	Sequence 8, Appl1
11	199.5	68.8	40	1	PCT-US00-22121-35	Sequence 35, Appl1
12	199.5	68.8	40	14	US-09-026-276-35	Sequence 35, Appl1
13	199.5	68.8	40	17	US-09-374-721A-35	Sequence 35, Appl1
14	199.5	68.8	41	1	PCT-US00-22121-34	Sequence 34, Appl1
15	199.5	68.8	41	14	US-09-026-276-34	Sequence 34, Appl1
16	199.5	68.8	41	17	US-09-374-721A-34	Sequence 34, Appl1
17	187.5	64.7	40	19	US-09-506-078-15	Sequence 15, Appl1
18	187.5	64.7	40	22	US-60-120-454-15	Sequence 15, Appl1
19	187.5	64.7	44	5	US-08-160-882-45	Sequence 45, Appl1
20	187.5	64.7	84	5	US-08-160-882-47	Sequence 47, Appl1
21	187.5	64.7	397	19	US-09-506-078-31	Sequence 31, Appl1
22	187.5	64.7	397	22	US-60-120-454-31	Sequence 31, Appl1
23	187.5	64.7	398	19	US-09-506-078-25	Sequence 25, Appl1
24	187.5	64.7	398	22	US-60-120-454-25	Sequence 25, Appl1
25	187.5	64.7	411	19	US-09-506-078-23	Sequence 23, Appl1
26	187.5	64.7	411	22	US-60-120-454-23	Sequence 23, Appl1
27	187.5	64.7	442	19	US-09-506-078-27	Sequence 27, Appl1
28	187.5	64.7	442	22	US-60-120-454-27	Sequence 27, Appl1
29	108	37.2	20	1	PCT-US00-22121-26	Sequence 26, Appl1
30	108	37.2	20	8	US-08-476-013-12	Sequence 12, Appl1
31	108	37.2	20	14	US-09-026-276-26	Sequence 26, Appl1
32	108	37.2	20	17	US-09-374-721A-26	Sequence 26, Appl1
33	108	37.2	21	8	US-08-477-298-14	Sequence 14, Appl1
34	106	36.6	20	3	US-07-761-849-10	Sequence 10, Appl1
35	106	36.6	20	8	US-08-476-013-10	Sequence 10, Appl1
36	105	36.2	20	1	PCT-US00-22121-30	Sequence 30, Appl1
37	105	36.2	20	1	PCT-US00-22121-31	Sequence 31, Appl1
38	105	36.2	20	14	US-09-026-276-30	Sequence 30, Appl1
39	105	36.2	20	14	US-09-026-276-31	Sequence 31, Appl1
40	105	36.2	20	17	US-09-374-721A-30	Sequence 30, Appl1
41	105	36.2	21	3	US-09-374-721A-31	Sequence 31, Appl1
42	105	36.2	21	3	US-07-761-849-11	Sequence 11, Appl1
43	105	36.2	21	3	US-07-761-849-12	Sequence 12, Appl1
44	105	36.2	21	8	US-08-476-013-11	Sequence 11, Appl1
45	105	36.2	21	8	US-08-476-013-12	Sequence 12, Appl1

ALIGNMENTS

RESULT 1
US-09-019-010-4
: Sequence 4, Application US/09019010
: GENERAL INFORMATION:
: APPLICANT: HARLAND, RICHARD
: APPLICANT: MANN, JOHN G.
: APPLICANT: ACRES, STEPHEN D.
: TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS
: MOLECULES
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: ROBINS & ASSOCIATES
: STREET: 90 MIDDLEFIELD ROAD, SUITE 200
: CITY: MENLO PARK
: STATE: CA
: COUNTRY: USA
: ZIP: 94025
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/019,010
: FILING DATE: 05-FEB-1998
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/036,883

FILING DATE: 05-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0035
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 325-7812
TELEFAX: (650) 325-7823
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-019-010-4

Query Match 100.0%; Score 290; DB 14; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.7e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49
DB 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49

RESULT 2
US-09-249-447A-7
Sequence 7, Application US/09249447A
GENERAL INFORMATION:
APPLICANT: Manns, Jack G.
TITLE OF INVENTION: PASSIVE IMMUNIZATION AS A TREATMENT FOR HORMONE
FILE REFERENCE: 9001-0045
CURRENT APPLICATION NUMBER: US/09/249,447A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/075,637
PRIOR FILING DATE: 1998-02-23
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 49
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: GnRH-2, Fig.
US-09-249-447A-7

Query Match 100.0%; Score 290; DB 16; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.7e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49
DB 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49

RESULT 3
US-09-305-924-11
Sequence 11, Application US/09305924A
GENERAL INFORMATION:
APPLICANT: Jack G. Manns
APPLICANT: Stephen D. Acres
APPLICANT: Richard Harland
TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
FILE REFERENCE: 9001-0048
CURRENT APPLICATION NUMBER: US/09/305,924A
CURRENT FILING DATE: 1999-05-05
EARLIER APPLICATION NUMBER: US 60/084,217
EARLIER FILING DATE: 1998-05-05
NUMBER OF SEQ ID NOS: 14

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 49
TYPE: PRT
ORGANISM: GnRH
US-09-305-924-11

Query Match 100.0%; Score 290; DB 17; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.7e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49
DB 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49

RESULT 4
US-09-306-689-11
Sequence 11, Application US/09306689B
GENERAL INFORMATION:
APPLICANT: Robbins, Sarah C.
TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
FILE REFERENCE: 9001-0047
CURRENT APPLICATION NUMBER: US/09/306,689B
CURRENT FILING DATE: 1999-05-06
EARLIER APPLICATION NUMBER: US 60/088,024
EARLIER FILING DATE: 1998-06-04
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 49
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Sequence
US-09-306-689-11

Query Match 100.0%; Score 290; DB 17; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.7e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49
DB 1 QHWSYGLRPGSGSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49

RESULT 5
US-09-383-912-4
Sequence 4, Application US/09383912
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANNs, JOHN G.
TITLE OF INVENTION: GnRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/383,912
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-383-912-4

Query Match 100.0%; Score 290; DB 17; Length 49;
Best Local Similarity 100.0%; Pred. No. 4.7e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 49
|||||
Db 1 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 49

RESULT 6
US-09-383-912-10
Sequence 10, Application US/09383912
GENERAL INFORMATION:
APPLICANT: POTTER, ANDREW A.
APPLICANT: MANN, JOHN G.
TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: REED & ROBINS LLP
STREET: 285 HAMILTON AVENUE, SUITE 200
CITY: PALO ALTO
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/383,912
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/694,865
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 9001-0016.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)327-3400
TELEFAX: (415)327-3231
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 544 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-383-912-10

Query Match 100.0%; Score 290; DB 17; Length 544;

Best Local Similarity 100.0%; Pred. No. 5e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 49
|||||
Db 494 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 542

RESULT 7
US-09-305-924-13
Sequence 13, Application US/09305924A
GENERAL INFORMATION:
APPLICANT: Jack G. Manns
APPLICANT: Stephen D. Acres
TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
FILE REFERENCE: 9001-0048
CURRENT APPLICATION NUMBER: US/09/305,924A
CURRENT FILING DATE: 1999-05-05
EARLIER APPLICATION NUMBER: US 60/084,217
EARLIER FILING DATE: 1998-05-05
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 13
LENGTH: 695
TYPE: PRT
ORGANISM: GNRH
US-09-305-924-13

Query Match 100.0%; Score 290; DB 17; Length 695;
Best Local Similarity 100.0%; Pred. No. 6.4e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 49
|||||
Db 9 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 57

RESULT 8
US-09-306-689-13
Sequence 13, Application US/09306689B
GENERAL INFORMATION:
APPLICANT: Robbins, Sarah C.
TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
FILE REFERENCE: 9001-0047
CURRENT APPLICATION NUMBER: US/09/306,689B
CURRENT FILING DATE: 1999-05-06
EARLIER APPLICATION NUMBER: US 60/088,024
EARLIER FILING DATE: 1998-06-04
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 13
LENGTH: 695
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Sequence
US-09-306-689-13

Query Match 100.0%; Score 290; DB 17; Length 695;
Best Local Similarity 100.0%; Pred. No. 6.4e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 49
|||||
Db 9 QHWSYGLRPGSGSDMSYGLRPGSSQHWSYGLRPGSGSDMSYGLRPG 57

RESULT 9

US-09-383-912-16
; Sequence 16, Application US/09383912
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-383-912-16

Query Match 100.0%; Score 290; DB 17; Length 699;
Best Local Similarity 100.0%; Pred. No. 6.4e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49
DB 9 QHWSYGLRPGSSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 57

RESULT 10
US-09-383-912-8
; Sequence 8, Application US/09383912
; GENERAL INFORMATION:
; APPLICANT: POTTER, ANDREW A.
; TITLE OF INVENTION: GNRH-LEUKOTOXIN CHIMERAS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: REED & ROBINS LLP
; STREET: 285 HAMILTON AVENUE, SUITE 200
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,912

US-09-383-912-8
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,865
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCracken, THOMAS P.
; REGISTRATION NUMBER: 38,548
; REFERENCE/DOCKET NUMBER: 9001-0016.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)327-3400
; TELEFAX: (415)327-3231
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 977 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-383-912-8

Query Match 100.0%; Score 290; DB 17; Length 977;
Best Local Similarity 100.0%; Pred. No. 9e-23;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QHWSYGLRPGSSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49
DB 927 QHWSYGLRPGSSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 975

RESULT 11
PCT-US00-22121-35
; Sequence 35, Application PC/TUS0022121
; GENERAL INFORMATION:
; APPLICANT: Proteinix Company
; TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: IGN-2004WO
; CURRENT APPLICATION NUMBER: PCT/US00/22121
; CURRENT FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 09/026,276
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 09/374,721
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 35
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: polypeptide
; PCT-US00-22121-35

Query Match 68.8%; Score 199.5; DB 1; Length 40;
Best Local Similarity 77.6%; Pred. No. 1.6e-14;
Matches 38; Conservative 0; Mismatches 2; Indels 9; Gaps 3;

QY 1 QHWSYGLRPGSSQDWSYGLRPGSSQHWSYGLRPGSSQDWSYGLRPG 49
DB 1 QHWSYGLRPG---QHWSYGLRPG---QHWSYGLRPG---QHWSYGLRPG 40

RESULT 12
US-09-026-276-35
; Sequence 35, Application US/09026276
; GENERAL INFORMATION:
; APPLICANT: Kenten, John H
; APPLICANT: Tramontano, Alfonso
; APPLICANT: Pilon, April L
; APPLICANT: Lohmas, Gerald L
; APPLICANT: Roberts, Steven F

```

?
?
? TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
? FILE REFERENCE: U.S. Patent Application No. 09/026,276
? CURRENT APPLICATION NUMBER: US/09/026,276
? CURRENT FILING DATE: 1998-02-19
? NUMBER OF SEQ ID NOS: 35
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 35
?
? LENGTH: 40
?
? TYPE: PRT
?
? ORGANISM: Porcine
?
? US-09-026-276-35

```

Query Match	68.8%	Score 199.5	DB 14	Length 40
Best Local Similarity	77.6%	Pred. No. 1.6e14		
Matches 38	Conservative 0	Mismatches 2	Indels 9	Gaps 3

QY 1 QHWSYGLRPGSGSQDWSTYGLRPGSSQHWSYGLRPGSGSQDWSTYGLRPG 49
||||||| | ||||| | ||||| | ||||| |
Db 1 QHWSYGLRPG--QHWSYGLRPG--QHWSYGLRPG--QHWSYGLRPG 400

```

RESULT 13
US-09-374-721A-35
: Sequence 35, Application US/09374721A
: GENERAL INFORMATION:
: APPLICANT: Kenten, John H.
: APPLICANT: Roberts, Steven
: APPLICANT: Lohmas, Gerald
: TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
: FILE REFERENCE: CIP OF ION-6601
: CURRENT APPLICATION NUMBER: US/09/374,721A
: CURRENT FILING DATE: 1999-08-13
: PRIOR APPLICATION NUMBER: 09/006,276
: PRIOR FILING DATE: 1998-02-19
: NUMBER OF SEQ ID NOS: 35
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 35
: LENGTH: 40
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: polypeptide
: US-09-374-721A-35

```

Query Match	68.8%	Score 199.5:	DB 17:	Length 40:
Best Local Similarity	77.6%	Pred. NO. 1,6e-14:		
Matches 38:	Conservative 0:	Mismatches 2:	Indels 9:	Gaps 3:
OY	1	QHWSTGLRPGSGSQDMWSTGLRPGSSQHWSTGLRPGSGSQDMWSTGLRPG	49	
Db	1	QHWSTGLRPG---QHWSTGLRPG---QHWSTGLRPG---QHWSTGLRPG	40	

```

RESULT 14
PCT-US00-22121-34
Sequence 34, Application PC/TUS0022121
GENERAL INFORMATION:
APPLICANT: Proteinix Company
TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
FILE REFERENCE: IGN-2004WO
CURRENT APPLICATION NUMBER: PCT/US00/22121
PRIORITY FILING DATE: 2000-08-14
PRIORITY APPLICATION NUMBER: 09/026,276
PRIORITY FILING DATE: 1998-02-19
PRIORITY APPLICATION NUMBER: US 09/374,721
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 34
LENGTH: 41

```

```

: TYPE: PRT
:
: ORGANISM: Artificial Sequence
:
: FEATURE:
:
: OTHER INFORMATION: Description of Artificial Sequence: polypeptide
:
: OTHER INFORMATION: antigen
PCT-US00-22121-34

```

Query Match	68.8%;	Score 199.5;	DB 1;	Length 41;
Best Local Similarity	77.6%;	Pred. No. 1.6e-14;		
Matches 38;	Conservative 0;	Mismatches 2;	Indels 9;	Gaps 3;

```

OY 1 QHWSYGLRPGSGSQDWSTYGLRPGSSQHWSTYGLRPGSGSQDWSTYGLRPG 49
    |||||  | |||||  |||||  | |||||
DB 1 QHWSYGLRPG--QHWSYGLRPG--QHWSYGLRPG--QHWSYGLRPG 40

```

RESULT 15
US-09-026-276-34
Sequence 34, Application US/09026276
GENERAL INFORMATION:
APPLICANT: Kenten, John H
APPLICANT: Tiramontano, Alfonso
APPLICANT: Pilon, Aprille L
APPLICANT: Lohrns, Steven F
TITLE OF INVENTION: HEAT-SHOCK FUSION-BASED VACCINE SYSTEM
FILE REFERENCE: U.S. Patent Application No. 09\026,276
CURRENT APPLICATION NUMBER: US/09/026,276
CURRENT FILING DATE: 1998-02-19
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 34
LENGTH: 41
TYPE: PRT
ORGANISM: Porcine
US-09-026-276-34

	68.8%;	Score 199.5;	DB 14;	Length 41;
Best Local Similarity	77.6%;	Pred. No.	1.6e-14;	
Matches	38;	Conservative	0;	Mismatches 2; Indels 9; Gaps 3;
OY	1	QHWSTGLRPGSGSQDMWSTGLRPGSSQHMVSTGLRPGSSQDMWSTGLRPG	49	
DG	1	QHWSVTLRPG---QHWSVTLRPG---QHWSVTLRPG	40	

Search completed: March 2, 2001, 10:59:04
Job time: 374 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:44:35 ; Search time 47.48 Seconds
(without alignments)
5.295 Million cell updates/sec

Title: US-09-306-689-2
Perfect score: 58
Sequence: 1 XHMSYXLRPCXXXX 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents-AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/6.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/PCtUS.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/Backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	87.9	10	1	US-07-714-540-9
2	51	87.9	10	1	US-07-690-983D-2
3	51	87.9	10	1	US-07-690-983D-32
4	51	87.9	10	1	US-08-103-022-1
5	51	87.9	10	1	US-08-127-351-6
6	51	87.9	10	1	US-08-184-935-6
7	51	87.9	10	1	US-08-193-374-3
8	51	87.9	10	1	US-08-193-374-8
9	51	87.9	10	1	US-08-343-883-1
10	51	87.9	10	1	US-08-480-367B-6
11	51	87.9	10	1	US-08-000-931-5
12	51	87.9	10	1	US-08-487-221A-6
13	51	87.9	10	1	US-08-480-370-6
14	51	87.9	10	1	US-08-428-468-22
15	51	87.9	10	1	US-08-341-219-11
16	51	87.9	10	1	US-08-453-588-2
17	51	87.9	10	1	US-08-453-588-4
18	51	87.9	10	1	US-08-453-588-6
19	51	87.9	10	1	US-08-453-588-8
20	51	87.9	10	1	US-08-453-588-10
21	51	87.9	10	1	US-08-453-588-12
22	51	87.9	10	1	US-08-453-588-14
23	51	87.9	10	1	US-08-453-588-16
24	51	87.9	10	1	US-08-453-588-19
25	51	87.9	10	1	US-08-453-588-22
26	51	87.9	10	1	US-08-188-223-3
27	51	87.9	10	1	US-08-406-935-5
28	51	87.9	10	1	US-08-591-917-1

29	51	87.9	10	1	US-08-387-156-2	Sequence 2, Appli
30	51	87.9	10	1	US-08-474-555-1	Sequence 1, Appli
31	51	87.9	10	1	US-08-446-692-1	Sequence 1, Appli
32	51	87.9	10	1	US-08-242-678D-1	Sequence 1, Appli
33	51	87.9	10	1	US-08-242-678D-7	Sequence 6, Appli
34	51	87.9	10	2	US-08-796-598-6	Sequence 7, Appli
35	51	87.9	10	2	US-08-694-865-2	Sequence 2, Appli
36	51	87.9	10	2	US-08-694-865-18	Sequence 18, Appli
37	51	87.9	10	2	US-08-488-351A-1	Sequence 1, Appli
38	51	87.9	10	2	US-08-480-494B-1	Sequence 1, Appli
39	51	87.9	10	2	US-08-447-175A-6	Sequence 6, Appli
40	51	87.9	10	2	US-08-878-748-2	Sequence 2, Appli
41	51	87.9	10	3	US-08-521-079-2	Sequence 4, Appli
42	51	87.9	10	3	US-08-521-079-4	Sequence 4, Appli
43	51	87.9	10	3	US-08-521-079-6	Sequence 6, Appli
44	51	87.9	10	3	US-08-521-079-8	Sequence 8, Appli
45	51	87.9	10	5	5492893-1	Patent No. 5492893

ALIGNMENTS

```
RESULT 1
US-07-714-540-9
; Sequence 9, Application US/07714540
; Patent No. 5262521
; GENERAL INFORMATION:
; APPLICANT: Almqvist, Ronald G.
; APPLICANT: Toll, Lawrence
; TITLE OF INVENTION: ISOLATED ATRIAL PEPTIDE-DEGRADING
; TITLE OF INVENTION: ENZYME AND NOVEL COMPOUNDS USEFUL AS INHIBITORS THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESS: Irell & Manella
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07714,540
; FILING DATE: 19910607
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Dianne E.
; REGISTRATION NUMBER: 31,292
; REFERENCE/DOCKET NUMBER: 8500-0135.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-327-7250
; TELEFAX: 415-327-2951
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-714-540-9

Query Match      87.9%      Score 51: DB 1: Length 10;
Best Local Similarity 88.9%      Pred. No. 0.00095;
Matches      8: Conservative      0: Mismatches      1: Indels      0: Gaps      0:
QY      2 HMSYXLRPC 10
      |||||
DB      2 HMSYGLRPC 10
```

RESULT 2
US-07-690-983D-2
Sequence 2, Application US/07690983D
Patent No. 5403586
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J.
APPLICANT: STEWART, Andrew G.
APPLICANT: TSONIS, Con G.
TITLE OF INVENTION: FUSION PROTEINS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/690,983D
CLASSIFICATION: 435
FILING DATE: 25-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00373
FILING DATE: 24-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/148 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5390
TELEFAX: (202)672-5399
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-07-690-983D-2

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.00095;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
Db 2 HWSYGLRPG 10

RESULT 3
US-07-690-983D-32
Sequence 32, Application US/07690983D
Patent No. 5403586
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, Gregory J.
APPLICANT: STEWART, Andrew G.
APPLICANT: TSONIS, Con G.
TITLE OF INVENTION: FUSION PROTEINS
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/690,983D
CLASSIFICATION: 435
FILING DATE: 25-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU90/00373
FILING DATE: 24-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/148 CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5390
TELEFAX: (202)672-5399
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-690-983D-32

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.00095;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
Db 2 HWSYGLRPG 10

RESULT 4
US-08-103-022-1
Sequence 1, Application US/08103022
Patent No. 5413990
GENERAL INFORMATION:
APPLICANT: Haviv, Fortuna
APPLICANT: Fitzpatrick, Timothy D.
APPLICANT: Swenson, Rolf E.
APPLICANT: Nichols, Charles J.
APPLICANT: Mort, Nicholas A.
TITLE OF INVENTION: N-Terminus Modified Analogs of LHRH
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edward H. Gorman, Jr., Dept. 377
STREET: Abbott Laboratories, One Abbott Park Road
CITY: No. 5413990th Chicago
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/103,022
CLASSIFICATION: 514
FILING DATE: 05-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Janssen, Jerry F.
REGISTRATION NUMBER: 29,175
REFERENCE/DOCKET NUMBER: 5389.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 938-7742
TELEFAX: (708) 938-2623
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "Xaa at position 1 is a
OTHER INFORMATION: 5-oxo-prolyl aminoacyl residue."
US-08-103-022-1

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.00095;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
DB 2 HWSYGLRPG 10

RESULT 5
US-08-127-351-6
Sequence 6, Application US/08127351
Patent No. 5449761
GENERAL INFORMATION:
APPLICANT: BELINKA JR, BENJAMIN A.
APPLICANT: COUGHLIN, DANIEL J.
APPLICANT: ALVAREZ, VERNON L.
APPLICANT: WOOD, RICHARD
TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
TITLE OF INVENTION: CONSTRUCTS
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &
ADDRESSEE: NEUSTADT,
ADDRESS: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/127,351
FILING DATE: 28-SEP-1993
CLASSIFICATION: 534
ATTORNEY/AGENT INFORMATION:
NAME: Villacorta, Gilberto M.
REGISTRATION NUMBER: 34,038
REFERENCE/DOCKET NUMBER: 4980-004-44
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 6
OTHER INFORMATION: /note= "Position 6 may be either
OTHER INFORMATION: Gly or D-Trp."
US-08-127-351-6

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00095;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
DB 2 HWSYXLRPG 10

RESULT 6
US-08-184-935-6
Sequence 6, Application US/08184935
Patent No. 5476770
GENERAL INFORMATION:
APPLICANT: PRADELLES, PHILIPPE
TITLE OF INVENTION: IMMUNOMETRIC DETERMINATION OF AN ANTIGEN
TITLE OF INVENTION: OR HAPTEN
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/184,935
FILING DATE: 24-JAN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5476770man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 846-286-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 10
OTHER INFORMATION: /note= "C-terminal amide"
US-08-184-935-6

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.00095;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
DB 2 HWSYGLRPG 10

RESULT 7
US-08-193-374-3
Sequence 3, Application US/08193374
Patent No. 5487898
GENERAL INFORMATION:
APPLICANT: Fu Lu, Mou-Ying
APPLICANT: Reiland, Thomas L.
TITLE OF INVENTION: Compositions and Method for the

TITLE OF INVENTION: Sublingual or Buccal Administration of Therapeutic Agents
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edward H. Gorman, Jr., Dept. 377 - AP6D
STREET: Abbott Laboratories, One Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/193,374
FILING DATE: 07-FEB-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07 / 983,111
FILING DATE: 30-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Janssen, Jerry F.
REGISTRATION NUMBER: 29,175
REFERENCE/DOCKET NUMBER: 4848.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 938-7742
TELEFAX: (708) 938-2623
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note="XAA at position 1 is a
OTHER INFORMATION: 5-oxopropyl residue."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 6
OTHER INFORMATION: /note="XAA at position 6 is a
OTHER INFORMATION: D-tryptyl residue."
US-08-193-374-3

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00095;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
DB 2 HWSYXLRPG 10

RESULT 8
US-08-193-374-8
Sequence 8, Application US/08193374
Patent No. 5487898
GENERAL INFORMATION:
APPLICANT: Fu Lu, Mou-Ying
APPLICANT: Reiland, Thomas L.
TITLE OF INVENTION: Compositions and Method for the
TITLE OF INVENTION: Sublingual or Buccal Administration of Therapeutic Agents
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edward H. Gorman, Jr., Dept. 377 - AP6D
STREET: Abbott Laboratories, One Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA

ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/193,374
FILING DATE: 07-FEB-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07 / 983,111
FILING DATE: 30-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Janssen, Jerry F.
REGISTRATION NUMBER: 29,175
REFERENCE/DOCKET NUMBER: 4848.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 938-7742
TELEFAX: (708) 938-2623
INFORMATION FOR SEQ. ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note="XAA at position 1 is a
OTHER INFORMATION: 5-oxopropyl residue."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 6
OTHER INFORMATION: /note="XAA at position 6 is a
OTHER INFORMATION: D-3-(naphth-2-yl)alanyl residue."
US-08-193-374-8

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00095;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
DB 2 HWSYXLRPG 10

RESULT 9
US-08-343-883-1
Sequence 1, Application US/08343883
Patent No. 5573767
GENERAL INFORMATION:
APPLICANT: Dufour, Raymond J.
APPLICANT: Roulet, Claude J.M.
APPLICANT: Chouvet, Claire D.
APPLICANT: Bonneau, Michel B.
TITLE OF INVENTION: Method for improving the organoleptic
TITLE OF INVENTION: qualities of the meat from uncastrated male domestic
TITLE OF INVENTION: animals, vaccines which are usable in this method, new
TITLE OF INVENTION: peptide, in particular for producing these vaccines...
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Larson and Taylor
STREET: 727 Twenty-Third Street, South
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343,883
FILING DATE: 17-NOV-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/946,495
FILING DATE: 09-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9102513
FILING DATE: 01-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9115289
FILING DATE: 10-DEC-1991
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 10
OTHER INFORMATION: /label= NH2
OTHER INFORMATION: /note= "amidated glycine"
FEATURE:
NAME/KEY: Peptide
LOCATION: 1
OTHER INFORMATION: /label= pyro
OTHER INFORMATION: /note= "pyroglutamic acid"
PUBLICATION INFORMATION:
AUTHORS: Matsuo, H.
AUTHORS: Baba, Y.
AUTHORS: G. Nair, R. M.
AUTHORS: Arimura, A.
AUTHORS: Schally, A. V.
TITLE: Structure of the porcine LH- and
TITLE: FSH-releasing hormone. I. The proposed amino acid
TITLE: sequence.
JOURNAL: Biochem. Biophys. Res. Commun.
VOLUME: 43
ISSUE: 6
PAGES: 1334-1339
DATE: 1971
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 10
US-08-343-883-1

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.00095;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 HWSYXLRPG 10
| | | | |
Db 2 HWSYGLRPG 10

RESULT 10
US-08-480-367B-6
Sequence 6, Application US/08480367B
Patent No. 5578288
GENERAL INFORMATION:
APPLICANT: BELINKA JR, BENJAMIN A.
APPLICANT: COUGHLIN, DANIEL J.
APPLICANT: ALVAREZ, VERNON L.
APPLICANT: WOOD, RICHARD
TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
TITLE OF INVENTION: CONSTRUCTS
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER
STREET: 99 Canal Center Plaza, Suite 300

CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,367B
FILING DATE: 07-06-95
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Villacorta, Gilberto M.
REGISTRATION NUMBER: 34,038
REFERENCE/DOCKET NUMBER: 2654-002A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 684-1111
TELEFAX: (703) 684-1124
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 6
OTHER INFORMATION: /note= "Position 6 may be either
OTHER INFORMATION: Gly or D-Tip."
US-08-480-367B-6

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00095;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 HWSYXLRPG 10
| | | | |
Db 2 HWSYXLRPG 10

RESULT 11
US-08-000-931-5
Sequence 5, Application US/08000931
Patent No. 5578477
GENERAL INFORMATION:
APPLICANT: TAMANOI DR., FUYUHIKO
TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF
TITLE OF INVENTION: INHIBITORS OF PROTEIN FARNESYLTRANSFERASE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/000,931
FILING DATE: 05-JAN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 64098/102/ARDE
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-000-931-5

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.00095;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
DB 2 HWSYXLRPG 10

RESULT 12
US-08-487-221A-6
Sequence 6, Application US/08487221A
Patent No. 5593656
GENERAL INFORMATION:
APPLICANT: BELINKA JR, BENJAMIN A.
APPLICANT: COUGHLIN, DANIEL J.
APPLICANT: ALVAREZ, VERNON L.
APPLICANT: WOOD, RICHARD
TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
TITLE OF INVENTION: CONSTRUCTS
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER &
ADDRESS: NEUSTADT, P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,221A
CLASSIFICATION: 424
FILING DATE: 07-JUN-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/127,351
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Villacorta, Gilberto M.
REGISTRATION NUMBER: 34,038
REFERENCE/DOCKET NUMBER: 4980-004-44
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 6
OTHER INFORMATION: /note= "Position 6 may be either
OTHER INFORMATION: Gly or D-Tyr."

US-08-487-221A-6

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00095;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
DB 2 HWSYXLRPG 10

RESULT 13
US-08-480-370-6
Sequence 6, Application US/08480370
Patent No. 5609847
GENERAL INFORMATION:
APPLICANT: BELINKA JR, BENJAMIN A.
APPLICANT: COUGHLIN, DANIEL J.
APPLICANT: ALVAREZ, VERNON L.
APPLICANT: WOOD, RICHARD
TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
TITLE OF INVENTION: CONSTRUCTS
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER &
ADDRESS: NEUSTADT, P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,370
FILING DATE:
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/127,351
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Villacorta, Gilberto M.
REGISTRATION NUMBER: 34,038
REFERENCE/DOCKET NUMBER: 4980-004-44
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 6
OTHER INFORMATION: /note= "Position 6 may be either
OTHER INFORMATION: Gly or D-Tyr."

US-08-480-370-6

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00095;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
DB 2 HWSYXLRPG 10

Db 2 HWSYXLRPG 10

RESULT 14
US-08-428-488-22
; Sequence 22, Application US/08428488
; Patent No. 5624894
; GENERAL INFORMATION:
; APPLICANT: BODOR, Nicholas S.
; TITLE OF INVENTION: BRAIN-ENHANCED DELIVERY OF NEUROACTIVE
; TITLE OF INVENTION: PEPTIDES BY SEQUENTIAL METABOLISM
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P. O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/428.488
; FILING DATE: 27-APR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Baumeister, Mary Katherine
; REGISTRATION NUMBER: 26,254
; REFERENCE/DOCKET NUMBER: 028724-087
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "Position 1 = p-Glu."
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 10
; OTHER INFORMATION: /note= "Position 10 = Gly-NH2."
US-08-428-488-22

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.00095;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
|||||
Db 2 HWSYXLRPG 10

RESULT 15
US-08-341-219-11
; Sequence 11, Application US/08341219
; Patent No. 5643877
; GENERAL INFORMATION:
; APPLICANT: Zohar, Y.
; APPLICANT: Rivier, J.
; APPLICANT: Powell, J.
; APPLICANT: Sherwood, N.
; APPLICANT: Gothliff, Y.
; TITLE OF INVENTION: Compounds and Methods For Controlling

; TITLE OF INVENTION: Reproduction in Fish
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: N.Y.
; COUNTRY: USA
; ZIP: 10036-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/341.219
; FILING DATE: 05-DEC-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30742
; REFERENCE/DOCKET NUMBER: 8399-003-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /label= Glu1
; OTHER INFORMATION: /note= "pyroglutamic acid"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 10
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /label= Gly10
; OTHER INFORMATION: /note= "amidated"
US-08-341-219-11

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.00095;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
|||||
Db 2 HWSYXLRPG 10

Search completed: March 2, 2001, 10:54:36
Job time: 601 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:52:50 ; Search time 201.99 Seconds
(without alignments)
10.640 Million cell updates/sec

Title: US-09-306-689-2
Perfect score: 58
Sequence: 1 XHMSYXLRPCXXXX 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 957798 seqs, 153513742 residues
Total number of hits satisfying chosen parameters: 957798

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending-Patents-AA:*

1:	/cgn2_6/ptodata/2/paa/PCRTUS_COMB.pep.*
2:	/cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5:	/cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6:	/cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8:	/cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9:	/cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11:	/cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12:	/cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13:	/cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14:	/cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15:	/cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16:	/cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17:	/cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18:	/cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19:	/cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20:	/cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21:	/cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22:	/cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23:	/cgn2_6/ptodata/2/paa/PCRNEM_COMB.pep.*
24:	/cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
25:	/cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
26:	/cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
27:	/cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
28:	/cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	89.7	9	20	US-09-657-276-160
2	52	89.7	10	20	US-09-657-276-139
3	51	87.9	9	20	US-09-657-276-153
4	51	87.9	10	1	PCT-US00-22121-28
5	51	87.9	10	1	PCT-US00-22121-32

6	51	87.9	10	1	PCT-US94-04832A-1	Sequence 1, Appl
7	51	87.9	10	1	PCT-US94-12763-4	Sequence 4, Appl
8	51	87.9	10	1	PCT-US94-13394-5	Sequence 5, Appl
9	51	87.9	10	1	PCT-US95-01225-3	Sequence 3, Appl
10	51	87.9	10	1	PCT-US96-16950-2	Sequence 2, Appl
11	51	87.9	10	1	PCT-US96-16950-3	Sequence 3, Appl
12	51	87.9	10	1	PCT-US96-16950-7	Sequence 7, Appl
13	51	87.9	10	1	PCT-US96-16950-8	Sequence 8, Appl
14	51	87.9	10	1	PCT-US96-16950-9	Sequence 9, Appl
15	51	87.9	10	1	PCT-US96-17008-3	Sequence 3, Appl
16	51	87.9	10	1	PCT-US96-17008-7	Sequence 7, Appl
17	51	87.9	10	1	PCT-US96-17008-8	Sequence 8, Appl
18	51	87.9	10	1	PCT-US96-17008-9	Sequence 9, Appl
19	51	87.9	10	1	PCT-US96-17008-10	Sequence 10, Appl
20	51	87.9	10	1	PCT-US96-17008-13	Sequence 13, Appl
21	51	87.9	10	1	PCT-US99-11219-1143	Sequence 1143, Ap
22	51	87.9	10	1	PCT-US99-11219-1309	Sequence 1309, Ap
23	51	87.9	10	1	PCT-US99-11219-1344	Sequence 1344, Ap
24	51	87.9	10	1	PCT-US99-13960-28	Sequence 28, Appl
25	51	87.9	10	1	PCT-US99-13975B-77	Sequence 77, Appl
26	51	87.9	10	3	US-07-669-695-1	Sequence 1, Appl
27	51	87.9	10	3	US-07-672-300A-14	Sequence 14, Appl
28	51	87.9	10	3	US-07-728-782A-1	Sequence 1, Appl
29	51	87.9	10	3	US-07-728-782-1	Sequence 1, Appl
30	51	87.9	10	3	US-07-761-849-1	Sequence 1, Appl
31	51	87.9	10	3	US-07-946-062-22	Sequence 22, Appl
32	51	87.9	10	3	US-07-984-293-5	Sequence 5, Appl
33	51	87.9	10	4	US-08-020-366-9	Sequence 9, Appl
34	51	87.9	10	4	US-08-020-366-9	Sequence 9, Appl
35	51	87.9	10	4	US-08-057-166-1	Sequence 1, Appl
36	51	87.9	10	5	US-08-138-514-4	Sequence 4, Appl
37	51	87.9	10	5	US-08-138-514-11	Sequence 11, Appl
38	51	87.9	10	5	US-08-138-514-12	Sequence 12, Appl
39	51	87.9	10	5	US-08-138-514-16	Sequence 16, Appl
40	51	87.9	10	5	US-08-138-514-20	Sequence 20, Appl
41	51	87.9	10	5	US-08-138-516-1	Sequence 1, Appl
42	51	87.9	10	5	US-08-138-516-6	Sequence 6, Appl
43	51	87.9	10	5	US-08-138-516-7	Sequence 7, Appl
44	51	87.9	10	5	US-08-138-516-11	Sequence 11, Appl
45	51	87.9	10	27	US-09-412-558-1	Sequence 1, Appl

ALIGNMENTS

US-09-657-276-160
: Sequence 160, Application US/09657276
: GENERAL INFORMATION:
: APPLICANT: Conjuchem, Inc.
: APPLICANT: Bridon, Dominique
: APPLICANT: Ezrin, Alan
: APPLICANT: Milner, Peter
: APPLICANT: Holmes, Darren
: APPLICANT: Thibaudau, Karen
: TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM
: TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD
: FILE REFERENCE: 2110
: CURRENT APPLICATION NUMBER: US/09/657,276
: CURRENT FILING DATE: 2000-09-07
: PRIOR APPLICATION NUMBER: 60/134,406
: PRIOR FILING DATE: 1999-05-17
: PRIOR APPLICATION NUMBER: 60/153,406
: PRIOR FILING DATE: 1999-09-10
: PRIOR APPLICATION NUMBER: 60/159,783
: PRIOR FILING DATE: 1999-10-18
: NUMBER OF SEQ ID NOS: 1617
: SOFTWARE: Patent Ver. 2.1
: SEQ ID NO 160
: LENGTH: 9
: TYPE: PRT
: ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Peptide
US-09-657-276-160

Query Match 89.7%; Score 52; DB 20; Length 9;
Best Local Similarity 88.9%; Pred. No. 8.3e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
|||||
Db 1 HWSYSLRPG 9

RESULT 2
US-09-657-276-139
Sequence 139, Application US/09657276

GENERAL INFORMATION:
APPLICANT: Conjuchem, Inc.
APPLICANT: Bridon, Dominique
APPLICANT: Ezrin, Alan
APPLICANT: Milner, Peter
APPLICANT: Holmes, Darren
APPLICANT: Thibaudau, Karen
TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM
TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD
FILE REFERENCE: 2110
CURRENT APPLICATION NUMBER: US/05/657, 276
CURRENT FILING DATE: 2000-09-07
PRIOR APPLICATION NUMBER: 60/134, 406
PRIOR FILING DATE: 1999-05-17
PRIOR APPLICATION NUMBER: 60/153, 406
PRIOR FILING DATE: 1999-09-10
PRIOR APPLICATION NUMBER: 60/155, 783
PRIOR FILING DATE: 1999-10-18
NUMBER OF SEQ ID NOS: 1617
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 139
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-657-276-139

Query Match 89.7%; Score 52; DB 20; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.015;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
|||||
Db 2 HWSYALRPG 10

RESULT 3
US-09-657-276-153
Sequence 153, Application US/09657276
GENERAL INFORMATION:
APPLICANT: Conjuchem, Inc.
APPLICANT: Bridon, Dominique
APPLICANT: Ezrin, Alan
APPLICANT: Milner, Peter
APPLICANT: Holmes, Darren
APPLICANT: Thibaudau, Karen
TITLE OF INVENTION: PROTECTION OF ENDOGENOUS THERAPEUTIC PEPTIDES FROM
TITLE OF INVENTION: PEPTIDASE ACTIVITY THROUGH CONJUGATION TO BLOOD
FILE REFERENCE: 2110
CURRENT APPLICATION NUMBER: US/09/657, 276

CURRENT FILING DATE: 2000-09-07
PRIOR APPLICATION NUMBER: 60/134, 406
PRIOR FILING DATE: 1999-05-17
PRIOR APPLICATION NUMBER: 60/153, 406
PRIOR FILING DATE: 1999-09-10
PRIOR APPLICATION NUMBER: 60/159, 783
PRIOR FILING DATE: 1999-10-18
NUMBER OF SEQ ID NOS: 1617
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 153
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-657-276-153

Query Match 87.9%; Score 51; DB 20; Length 9;
Best Local Similarity 88.9%; Pred. No. 8.3e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
|||||
Db 1 HWSYGLRPG 9

RESULT 4
PCT-US00-22121-28
Sequence 28, Application PC/TUS0022121
GENERAL INFORMATION:
APPLICANT: Proteinix Company
TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
FILE REFERENCE: IGN-2004MO
CURRENT APPLICATION NUMBER: PCT/US00/22121
CURRENT FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 09/026, 276
PRIOR FILING DATE: 1998-02-19
PRIOR APPLICATION NUMBER: US 09/374, 721
PRIOR FILING DATE: 1999-08-13
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 28
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: polypeptide
PCT-US00-22121-28

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.023;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYXLRPG 10
|||||
Db 2 HWSYGLRPG 10

RESULT 5
PCT-US00-22121-32
Sequence 32, Application PC/TUS0022121
GENERAL INFORMATION:
APPLICANT: Proteinix Company
TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
FILE REFERENCE: IGN-2004MO
CURRENT APPLICATION NUMBER: PCT/US00/22121
CURRENT FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 09/026, 276
PRIOR FILING DATE: 1998-02-19

PRIOR APPLICATION NUMBER: US 09/374,721
PRIOR FILING DATE: 1999-08-13
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentl Ver. 2.0
SEQ ID NO 32
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: polypeptide
PCT-US00-22121-32

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.023;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
||| |||
DB 2 HWSYGLRPG 10

RESULT 6
PCT-US94-04832A-1
Sequence 1, Application PC/TUS9404832A
GENERAL INFORMATION:

APPLICANT: Ladd, Anna
APPLICANT: Wang, Chang Yi
APPLICANT: Zamb, Timothy
TITLE OF INVENTION: Immunogenic LHRH peptide constructs
TITLE OF INVENTION: and synthetic universal immune stimulatoers for vaccines
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE:
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: US
ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentl Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04832A
FILING DATE: 13-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516)742-4343
TELEFAX: (516)742-4366
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US94-04832A-1

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.023;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
||| |||
DB 2 HWSYGLRPG 10

RESULT 7
PCT-US94-12763-4
Sequence 4, Application PC/TUS9412763
GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: Nucleic Acid Encoding [His-5,Tyr-7,Tyr-8]-
TITLE OF INVENTION: GnRH Preprohormone and [Ser-6]-GnRH preprohormone and
TITLE OF INVENTION: Their Uses
NUMBER OF SEQUENCES: 31
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentl Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/12763
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/147,771
FILING DATE: 05-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 14210-000400PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Mammal

FEATURE: Peptide
NAME/KEY: 1..10
LOCATION: 1..10
OTHER INFORMATION: /note="GnRH"
PCT-US94-12763-4

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.023;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPG 10
||| |||
DB 2 HWSYGLRPG 10

RESULT 8
PCT-US94-13394-5
Sequence 5, Application PC/TUS9413394
GENERAL INFORMATION:

APPLICANT: Seang H. Yiv
TITLE OF INVENTION: Transparent Liquid for
TITLE OF INVENTION: Encapsulating Drug Delivery
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and
ADDRESS: Norris
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/13394
FILING DATE: herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 885,202
FILING DATE: May 20, 1992
ATTORNEY/AGENT INFORMATION:
NAME: David R. Bailey
REGISTRATION NUMBER: 35,057
REFERENCE/DOCKET NUMBER: AFB1-0349
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acid residues
TYPE: Amino Acid
STRANDEDNESS:
MOLECULE TYPE: Peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: pyroglutamic acid
PCT-US94-13394-5

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.023;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
|||||
Db 2 HWSYGLRPG 10

RESULT 9
PCT-US95-01225-3
Sequence 3, Application PC/TUS9501225
GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
APPLICANT: Scibienski, Robert
TITLE OF INVENTION: Immunogens Against Gonadotropin
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dimitrios T. Divas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01225
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Divas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-354-8113
TELEFAX: 212-354-8286
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: YES
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label= pglu
OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..10
OTHER INFORMATION: /note= "immunomimic"
PCT-US95-01225-3

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.023;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
|||||
Db 2 HWSYGLRPG 10

RESULT 10
PCT-US96-16950-2
Sequence 2, Application PC/TUS9616950
GENERAL INFORMATION:
APPLICANT: LOMBARDO, VICTORIA K.
APPLICANT: MARBURG, STEPHEN
APPLICANT: TOLMAN, RICHARD L.
TITLE OF INVENTION: CONJUGATES OF GONADOTROPIN RELEASING
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: MOLLIE M. YANG
STREET: 126 E. LINCOLN AVE PO BOX 2000
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/16950
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: YANG, MOLLIE M.
REGISTRATION NUMBER: 32,718
REFERENCE/DOCKET NUMBER: 19444
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-6343
TELEFAX: 908-594-4720
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
PCT-US96-16950-2

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.023;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
11111111
DB 2 HWSYXLRPG 10

RESULT 11

PCT-US96-16950-3

Sequence 3, Application PC/TUS9616950

GENERAL INFORMATION:

APPLICANT: LOMBARDO, VICTORIA K.

APPLICANT: MARBURG, STEPHEN

APPLICANT: TOLMAN, RICHARD L.

TITLE OF INVENTION: CONJUGATES OF GONADOTROPIN RELEASING

TITLE OF INVENTION: HORMONE

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: MOLLIE M. YANG

STREET: 126 E. LINCOLN AVE PO BOX 2000

CITY: RAHWAY

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07065

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/16950

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: YANG, MOLLIE M.

REGISTRATION NUMBER: 32,718

REFERENCE/DOCKET NUMBER: 19444

TELECOMMUNICATION INFORMATION:

TELEPHONE: 908-594-6343

TELEFAX: 908-594-4720

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

PCT-US96-16950-3

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 HWSYXLRPG 10
11111111
DB 2 HWSYXLRPG 10

RESULT 12

PCT-US96-16950-7

Sequence 7, Application PC/TUS9616950

GENERAL INFORMATION:

APPLICANT: LOMBARDO, VICTORIA K.

APPLICANT: MARBURG, STEPHEN

APPLICANT: TOLMAN, RICHARD L.

TITLE OF INVENTION: CONJUGATES OF GONADOTROPIN RELEASING

TITLE OF INVENTION: HORMONE

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: MOLLIE M. YANG

STREET: 126 E. LINCOLN AVE PO BOX 2000

CITY: RAHWAY

STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/16950
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: YANG, MOLLIE M.
REGISTRATION NUMBER: 32,718
REFERENCE/DOCKET NUMBER: 19444
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-6343
TELEFAX: 908-594-4720
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
PCT-US96-16950-7

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 HWSYXLRPG 10
11111111
DB 2 HWSYXLRPG 10

RESULT 13

PCT-US96-16950-8

Sequence 8, Application PC/TUS9616950

GENERAL INFORMATION:

APPLICANT: LOMBARDO, VICTORIA K.

APPLICANT: MARBURG, STEPHEN

APPLICANT: TOLMAN, RICHARD L.

TITLE OF INVENTION: CONJUGATES OF GONADOTROPIN RELEASING

TITLE OF INVENTION: HORMONE

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: MOLLIE M. YANG

STREET: 126 E. LINCOLN AVE PO BOX 2000

CITY: RAHWAY

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07065

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/16950

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: YANG, MOLLIE M.

REGISTRATION NUMBER: 32,718

REFERENCE/DOCKET NUMBER: 19444

TELECOMMUNICATION INFORMATION:

TELEPHONE: 908-594-6343

TELEFAX: 908-594-4720

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
PCT-US96-16950-8

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.023;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
|||
Db 2 HWSYXLRPG 10

RESULT 14
PCT-US96-16950-9
Sequence 9, Application PC/TUS9616950
GENERAL INFORMATION:
APPLICANT: LOMBARDO, VICTORIA K.
APPLICANT: MARBURG, STEPHEN
APPLICANT: TOLMAN, RICHARD L.
TITLE OF INVENTION: CONJUGATES OF GONADOTROPIN RELEASING
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: MOLLIE M. YANG
STREET: 126 E. LINCOLN AVE PO BOX 2000
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/16950
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: YANG, MOLLIE M.
REGISTRATION NUMBER: 32,718
REFERENCE/DOCKET NUMBER: 19444
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-6343
TELEFAX: 908-594-4720
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
PCT-US96-16950-9

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
|||||
Db 2 HWSYXLRPG 10

RESULT 15

PCT-US96-17008-3
Sequence 3, Application PC/TUS9617008
GENERAL INFORMATION:
APPLICANT: Hickey, Gerard J.
APPLICANT: Mohr, Kenneth L.
TITLE OF INVENTION: PSEUDOMONAS EXOTOXIN AS IMMUNOGENIC
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: MOLLIE M. YANG
STREET: 126 E. Lincoln Ave. P.O. Box 2000
CITY: RAHWAY
STATE: New Jersey
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/17008
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Yang, Mollie M.
REGISTRATION NUMBER: 32,718
REFERENCE/DOCKET NUMBER: 19445
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-594-6343
TELEFAX: 908-594-4720
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US96-17008-3

Query Match 87.9%; Score 51; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYXLRPG 10
|||||
Db 2 HWSYXLRPG 10

Search completed: March 2, 2001, 10:59:02
Job time: 372 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:53:45 ; Search time 66.16 Seconds
(without alignments)
8.786 Million cell updates/sec

Title: US-09-306-689-3

Perfect score: 108

Sequence: 1 CPPPSEHMSYGLRPG 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.GeneSeq_36:*

1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	100.0	17	16	R78282
2	108	100.0	17	21	R58139
3	88.5	81.9	16	16	R78284
4	81	75.0	34	21	Y91244
5	81	75.0	34	21	Y68594
6	81	75.0	35	21	Y91242
7	81	75.0	35	21	Y91243
8	81	75.0	35	21	Y68593
9	71	65.7	49	17	W03944
10	71	65.7	49	19	W79567
11	71	65.7	49	19	W61542
12	71	65.7	49	21	Y58363

13	71	65.7	49	21	Y58135
14	71	65.7	544	17	W03943
15	71	65.7	544	19	W79570
16	71	65.7	695	19	W79573
17	71	65.7	695	21	Y58361
18	71	65.7	695	21	Y58361
19	71	65.7	977	17	W03942
20	71	65.7	977	19	W79569
21	69	63.9	92	7	P61427
22	68	63.0	253	12	R11181
23	68	63.0	263	12	R11185
24	68	63.0	323	12	R11187
25	66	61.1	28	15	R62698
26	66	61.1	256	12	R11177
27	65	60.2	29	16	R74262
28	65	60.2	253	12	R11184
29	64	59.3	11	19	W69276
30	64	59.3	16	15	R44956
31	64	59.3	936	14	R34547
32	63	58.3	10	2	P10411
33	63	58.3	10	2	P10416
34	63	58.3	10	6	P50222
35	63	58.3	10	7	P60127
36	63	58.3	10	7	P61403
37	63	58.3	10	7	P60576
38	63	58.3	10	8	P70922
39	63	58.3	10	10	P90630
40	63	58.3	10	12	R15713
41	63	58.3	10	13	R26819
42	63	58.3	10	15	R62689
43	63	58.3	10	16	R91197
44	63	58.3	10	16	R68645
45	63	58.3	10	21	Y53061

ALIGNMENTS

RESULT 1	
R78282	R78282 standard; peptide; 17 AA.
XX	
AC	R78282;
XX	
DT	13-MAR-1996 (first entry)
XX	
DE	GnRH immunomimetic and spacer (GnRH(1-10)-Ser1).
XX	
KW	Immunomimetic; gonadotropin releasing hormone; GnRH; LHRH; vaccine;
KW	luteinizing hormone releasing hormone; spacer; immunomimetic; uterine;
KW	diphtheria toxin; DT; gynaecological; endometriosis; uterine fibroids;
KW	gonadal steroid hormone associated dependent disease; gonadotropin;
KW	immunological contraception; mammal; breast; cancer; prostate;
KW	benign prostatic hypertrophy.
XX	
OS	Chimeric - Homo sapiens.
OS	Chimeric - Synthetic.
XX	
PH	Key
FT	Peptide
FT	Peptide
FT	Peptide
FT	Modified-site
XX	
PN	W09520600-A1.
XX	
PD	03-AUG-1995.
XX	
PF	26-JAN-1995; 95WO-US01225.
XX	
PR	27-JAN-1994; 94US-0188223.

GnRH analogue mult
LKT-GnRH protein f
LKT-GnRH chimeric
LKT-GnRH chimeric
Leukotoxin/gonadot
Gonadotropin relea
LKT-GnRH protein f
LKT-GnRH chimeric
Human preproLHRH p
Plasmid pBTA733-en
Plasmid pBTA870-en
Plasmid pBTA859-en
LHRH-containing im
Plasmid pBTA732-en
SSAL2 TH2LHRH. SY
Plasmid pBTA730-en
Luteinizing hormo
E. coli fimbriae H
GnRH-leukotoxin ge
Luteinizing Hormon
Luteinizing Hormon
Gonadotropin rele
Gonadoliberin anta
Gonadotropin relea
Novel decapeptide
Luteinizing Hormon
Sequence of lutein
Peptide #1 with ho
LH releasing hormo
LHRH happen for at
LHRH peptide. Syn
Gonadotropin relea
Luteinizing hormon

XX	(APHT-) APHTON CORP.
PA	
XX	
PI	Grimes S, Sciabienski R;
DR	WPI; 1995-275410/36.
XX	
XX	
PT	New anti-gonadotropin releasing hormone immunogenic composition(s)
PT	- used for treating gonadotropin and gonadal steroid hormone
PT	dependent disease(s) and providing contraception.
XX	
PS	Claim 1; Page 29; 39pp; English.
XX	
XX	The sequences given in R78282-85 represent immunomimetics to
CC	gonadotropin releasing hormone (GNRH). These peptides comprise the
CC	wild type GNRH sequence and a spacer attached to either the N- or C-
CC	terminal. The spacer molecules serve as a link through which the
CC	immunomimic is attached to an immunological carrier such as
CC	diphtheria toxoid (DT) and also affects the immune response generated
CC	by the vaccinated mammal against the immunomimic. Compositions
CC	comprising these peptides may be used for treating a mammal for
CC	gonadotropin and gonadal steroid hormone associated dependent
CC	disease or for providing immunological contraception in mammals.
CC	CC can also be used for treating breast cancer, uterine and other
CC	gynaecological cancers, endometriosis, uterine fibroids, prostate
CC	cancer, or benign prostatic hypertrophy.
XX	
XX	
Sequence	17 AA;
50	

Query Match	100.0%	Score 108;	DB 16	length 17;
Best Local Similarity	100.0%	Pred. No. 7.4e-08;		
Matches 17; Conservative	0;	Mismatches	0;	Gaps 0

QY	1	CPPPSSSEHWSYGLRPG	17
Db	1	CPPPPSSEHWSYGLRPG	17

RESULT 2

ID	standard; peptide; 17 AA.
Y58139	

AC Y58139;

DT 07-MAR-2000 (first entry)

DE Gonadotropin releasing hormone (GnRH), peptide analogue 3.

KM gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein
KM antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;
KM non-androgenic; steroid; reduction; weight gain; muscle distribution;
KM fat distribution; male pattern; boar taint; flavour; impalment;
KM reliable; immunocastration; meat production.

OS Synthetic.

PN W09956771-A2.

PD 11-NOV-1999

PF 05-MAY-1999; 99WO-CA00360.

PR 05-MAY-1998; 98US-0084217.

PA (BIOS-) BIOSTAR INC

PI Manns JG, Acres SD, Harland R;

DR DPT: 2000-062125/05.

PT Production of uncastrated male food animals using vaccines -

PS Disclosure; Page 11; 87pp; English
xx

CC Sequences Y58136-Y58141 represent gonadotropin releasing hormone
CC (GnRH) analogues which may be used as an alternative to sequence
CC Y58135 in embodiments of the present invention. The invention
CC relates to a method of using two GnRH immunogen vaccines to produce
CC uncastrated male animals for meat production, one vaccination prior to
CC or during the fattening period to reduce circulating testosterone levels
CC and the second vaccination about 2-8 weeks before slaughter to
CC substantially reduce androgenic and/or non-androgenic steroids. The
CC invention is used to produce food animals that exhibit the weight gain
CC and muscle/fat distribution of male animals without the problems
CC associated with male animals. Such problems include 'poor taint', a
CC urine-like odour found in cooked meat of uncastrated pigs which is
CC caused by steroids stored in the tissues, and similar flavour
CC impairments in the meat of other intact male animals. The invention is
CC more reliable than prior art immunocastration techniques.

SQ Sequence 17 AA;

Query Match	100.0%	Score 108; DB 21	Length 17;
Best Local Similarity	100.0%	Pred. No. 7.4e-08;	
Matches 17; Conservative	0;	Mismatches 0;	Indels 0; Gaps 0

```
QY 1 CPPPSSEHWSYGLRPG 17
    |||||
Db 1 cpppssehwsyglrpg 17
```

RESULT	3
R78284	
ID.	R78284 standard; peptide; 16 AA

AC R78284;

DT 13-MAR-1996 (first entry)

DE GnRH immunomimic and spacer (GnRH(1-10)-Arg1).

KW Immunomimetic; gonadotropin releasing hormone; GnRH; LHRH; vaccine;
KW luteinizing hormone releasing hormone; spacer; immunomimetic; uterine;
KW diptheria toxin; Dr; gynaecological; endometriosis; uterine fibroids
KW gonadal steroid hormone associated dependent disease; gonadotropin;
KW immunological contraceptive; mammal; breast; cancer; prostate;
KW benign prostatic hypertrophy.

OS Chimeric - Homo sapiens

XX

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

/note="spacer"

/note= "GnRH"

/note= "Amidated C-terminal"

PN MO9520600-AL
XX

03-AUG-1995

26-JAN-1995; 95WO-US01225.

PR 27-JAN-1994; 9405-0188223.

PA (ARHT-) APHTON CORP.

PI Grimes S, Scibienski R,

DR WPI: 1995-275410/36.

PT New anti-gonadotropin releasing hormone immunogenic composition(s),

PI - used for treating gonadotropin and gonadal steroid hormone
PR dependent disease(s) and providing contraception.
PS Claim 1; Page 29; 39pp; English.
XX The sequences given in R78282-85 represent immunomimetics to
CC gonadotropin releasing hormone (GnRH). These peptides comprise the
CC wild type GnRH sequence and a spacer attached to either the N- or C-
CC terminal. The spacer molecules serve as a link through which the
CC immunomimic is attached to an immunological carrier such as
CC diphtheria toxoid (DT) and also affects the immune response generated
CC by the vaccinated mammal against the immunomimic. Compositions
CC comprising these peptides may be used for treating a mammal for
CC gonadotropin and gonadal steroid hormone associated dependent
CC disease or for providing immunological contraception in mammals.
CC They can also be used for treating breast cancer, uterine and other
CC gynaecological cancers, endometriosis, uterine fibroids, prostate
CC cancer, or benign prostatic hypertrophy.
XX Sequence 16 AA:
SQ
Query Match 81.9%; Score 88.5; DB 16; Length 16;
Best Local Similarity 88.2%; Pred. No. 2.1e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
OY 1 CPPPSSEHWSYGLRPG 17
| | | | | | | | | | | | | | | | | |
DB 1 CPPPP-rehwsygltrpg 16
RESULT 4
Y91244
ID Y91244 standard; peptide: 34 AA.
XX Y91244;
XX 22-MAY-2000 (first entry)
XX Modified HBV surface Ag/LHRH antigenic peptide, SEQ ID NO:122.
DE Promiscuous T-cell epitope; measles virus F protein; MVF;
XX hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KM luteinizing hormone releasing hormone; LHRH; contraceptive; anticancer;
KM somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KM foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KM Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
KM cholesterol ester transport protein; anti-arteriosclerotic.
XX Chimeric - Hepatitis B virus.
OS WO966957-A2.
PN 29-DEC-1999.
PD 21-JUN-1999; 99WO-US13975.
PF 20-JUN-1998; 98US-0100412.
PR (UNBI-) UNITED BIOMEDICAL INC.
PA Wang CY.
XX WPI: 2000-160564/14.
DR
XX New artificial T helper cell epitope and derived immunogens with target
PR antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus
XX Disclosure: Page 109; 129pp; English.
XX The invention relates to novel promiscuous T helper cell epitopes (7h),
CC and immunogenic peptides comprising the 7h epitopes of the invention

CC along with B cell epitopes. The 7h epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response.
CC specifically against plasmodium falciparum, cholesterol ester transport
CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The 7h epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinizing hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target antigen.
CC Th can replace carrier proteins and pathogen-derived T helper epitopes.
CC Sequence Y91121 represents a promiscuous T helper epitope from the
CC measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and
CC Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th
CC epitope. Sequence Y91143 represents a promiscuous Th epitope from
CC hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are
CC synthetic epitopes derived from this HBV epitope. Y91156-Y91196,
CC Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH
CC sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target
CC antigenic peptide used in these LHRH antigenic peptides. Y91200 is
CC somatostatin, and Y91201-Y91207 are antigenic peptides comprising
CC somatostatin and a Th epitope. Somatostatin immunogens may be used
CC to promote growth in livestock. Y91208 is a human CD4 CDR-1-like domain
CC antigenic site, and Y91209-Y90211 are MWH Th epitope/CD4 CDR2
CC antigenic peptides which may be used to prevent HIV infection of T
CC cells. Y90212 is a modified version of a human IGE (immunoglobulin
CC E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic
CC peptides which may be used in the treatment of allergies. Y91220 is
CC a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid
CC protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223
CC is a Plasmodium falciparum circumsporozoite (CS) target antigen, and
CC Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be
CC used in a malaria vaccine. Y91228-Y91231 represent CERP-derived peptides
CC and Y91232-Y91241 are immunogens comprising a CERP peptide and a Th
CC epitope which may be used to prevent or treat arteriosclerosis and
CC cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising
CC B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic
CC peptides comprising MWH Th and HIV-1 B-cell epitope which may be used as
CC a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively
CC an immunostimulatory invasiv protein epitope from Yersinia species, and
CC hinge spacer peptide, both of which may optionally be used in the
CC antigenic peptides of the invention.
XX Sequence 34 AA:
SQ
Query Match 75.0%; Score 81; DB 21; Length 34;
Best Local Similarity 81.2%; Pred. No. 0.00038;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 2 PPPPSSEHWSYGLRPG 17
| | | | | | | | | | | | | | | | | |
DB 19 ppxpxpwhwsygltrpg 34
RESULT 5
Y68594
ID Y68594 standard; peptide: 34 AA.
XX Y68594;
XX 05-MAY-2000 (first entry)
XX Peptide immunogen comprising a Th epitope and LHRH target antigen.
XX Helper T cell epitope; surface antigen; peptide immunogen; LHRH;
KM luteinizing hormone-releasing hormone; spermatogenesis; ovulation;
KM oestrus; sexual development; sex hormone; promiscuous T helper epitope;

xx		Synthetic.		
OS	Hepatitis B virus.			
OS	Unidentified.			
xx				
FH	Key	Location/Qualifiers		
FT	Peptide	1..18		
FT		/note= "helper Th epitope"		
FT	Peptide	19..24		
FT		/note= "spacer Y68564"		
FT	Misc-difference	21		
FT		/note= "not specified"		
FT	Misc-difference	23		
FT		/note= "not specified"		
FT	Peptide	25..34		
FT		/note= "LHRH antigenic epitope Y68566"		
PN				
PD	W09966952-A1.			
xx				
xx	29-DEC-1999.			
PF				
PR	21-JUN-1999;	99WO-US13960.		
xx				
PR	20-JUN-1998;	98US-0100414.		
xx				
PA	(UNBI-) UNITED BIOMEDICAL INC.			
xx				
PI	Wang CY;			
DR				
WP1:	2000-160562/14.			
PT	New peptide immunogen containing luteinising hormone-releasing hormone			
PT	antigen site and helper T cell epitope, for e.g. contraceptive and			
PT	treatment of cancer -			
PS				
PS	Example 1; Page 85; 102pp; English.			
xx				
CC	The present sequence represents a peptide immunogen comprising a			
CC	synthetic helper T cell (Th) epitope and a target antigen, luteinising			
CC	hormone-releasing hormone (LHRH). The synthetic Th epitope is derived			
CC	from a structured synthetic antigen library (SSAL) designated SSAL2 Th2.			
CC	SSAL2 Th2 is modelled after a promiscuous epitope taken from the			
CC	Hepatitis B virus surface antigen. The peptide immunogens cause			
CC	induction of a specific immune response to LHRH which is involved in			
CC	regulation of spermatogenesis, ovulation, oestrus, sexual development			
CC	and secretion of sex hormones. Provision of a promiscuous T helper			
CC	epitope (which is functional in genetically diverse subjects) provides			
CC	optimum immunogenicity to the B cell epitopes of the target antigen and			
CC	thus high antibody titres against the target antigen. The peptide			
CC	immunogens of the invention are used to vaccinate against mammalian LHRH,			
CC	for use as (reversible) contraceptive; control of hormone-dependent			
CC	tumours (cancer of prostate or breast, also endometriosis); to prevent			
CC	boar taint (and improve meat quality) and for immunocastration.			
xx				
SQ	Sequence 34 AA;			
Query Match		75.0%;	Score 81;	DB 21; Length 34;
Best Local Similarity		81.2%;	Pred. No. 0.00038;	
Matches 13; Conservative		0;	Mismatches 3;	Indels 0; Gaps 0;
OY	2 PPPPSSEHSYGLRPG 17			
Db	19 PXPXPEHWSYGLRPg 34			
RESULT	6			
ID	Y91242			
XX	Y91242 standard; peptide: 35 AA.			
NC	Y91242;			

22-MAY-2000 (first entry)

Modified MZF Th epitope/LHRH antigenic peptide, SEQ ID NO:120.

Promiscuous T-cell epitope; measles virus F protein; MZF; hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope; interleukin hormone releasing hormone; LHRH; contraceptive; anticancer; somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV; foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic; Plasmodium falciparum; circumsporozoite; antimalarial; CERP; cholesterol ester transport protein; anti-arteriosclerotic.

Chimeric - Measles virus.

WO996957-A2.

29-DEC-1999.

21-JUN-1999; 99WO-US13975.

20-JUN-1998; 98US-0100412.

(UNB1-) UNITED BIOMEDICAL INC.

Wang CY;

WPI: 2000-160564/14.

New artificial T helper cell epitope and derived immunogens with target antigenic site, for immunization against e.g. malaria, arteriosclerosis or human immune deficiency virus -

Disclosure: Page 108; 129pp; English.

The invention relates to novel promiscuous T helper cell epitopes (Th), and immunogenic peptides comprising the Th epitopes of the invention along with B cell epitopes. The Th epitopes and peptide immunogens containing them, are used to induce a T helper cell response, specifically against Plasmodium falciparum, cholesterol ester transport protein (CERP) or HIV epitopes, but more generally against any pathogen, immunoreactive self-antigen or tumour antigen. The Th epitopes and peptide immunogens may be used for prevention and/or treatment of infections (HIV, foot-and-mouth disease or malaria); for cancer immunotherapy; for inhibition of the action of interleukin hormone releasing hormone (LHRH) for contraception, treatment of hormone-dependent cancer, prevention of boar taint in meat, and immunocastration; for promoting the growth of animals; or for treating allergies or arteriosclerosis. Incorporation of a promiscuous Th (functional in genetically diverse subjects) into an immunogen improves capacity to induce a strong T helper cell-mediated immune response, resulting in production of antibodies against a target antigen. Th can replace carrier proteins and pathogen-derived T helper epitopes. Sequence Y91121 represents a promiscuous T helper epitope from the measles virus F (MZF) protein and sequences Y91122-Y91142, Y91226 and Y91245-Y91246 represent synthetic Th epitopes based on the MZF Th epitope. Sequence Y91143 represents a promiscuous Th epitope from hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are synthetic epitopes derived from this HBV epitope. Y91156-Y91196, Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target antigenic peptide used in these LHRH antigenic peptides. Y91200 is somatostatin, and Y91201-Y91207 are antigenic peptides comprising somatostatin and a Th epitope. Somatostatin immunogens may be used to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain antigenic site, and Y91209-Y90211 are MZF Th epitopes/CD4 CDR2 antigenic peptides which may be used to prevent HIV infection of T cells. Y90212 is a modified version of a human IGE (immunoglobulin E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic peptides which may be used in the treatment of allergies. Y91220 is a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223 is a Plasmodium falciparum circumsporozoite (CS) target antigen, and

CC Y911224-Y911225 comprise the CS antigen and an MVE Th epitope and may be
CC used in a malaria vaccine. Y911228-Y911231 represent CERP-derived peptides
CC and Y911232-Y911241 are immunogens comprising a CERP peptide and a Th
CC epitope which may be used to prevent or treat arteriosclerosis and
CC cardiovascular disease. Y911247 and Y911252-Y911257 are HIV-1 neutralising
CC B-cell epitopes, and Y911248-Y911251 and Y911258-Y911273 are antigenic
CC peptides comprising MVA Th and HIV-1 B-cell epitope which may be used as
CC a component in an anti-HIV-1 vaccine. Y911198 and Y911199 are respectively
CC an immunostimulatory invasion protein epitope from *Yersinia* species, and
CC hinge spacer peptide, both of which may optionally be used in the
CC antigenic peptides of the invention.

50 **Sequence** **35** **AA;**

Query Match	75.0%	Score 81	DB 21	Length 35
Best Local Similarity	81.2%	Pred. No. 0.00039		
Matches 13, Conservative	0	Mismatches 3	Indels 0	Gaps 0

```
QY      2 PPPSSEHWSYGLRPG 17
          1111111111
Db      20 ppxpxpehwsyglrpg 35
```

RESULT	7
Y91243	
ID	Y91243 standard; peptide; 35 AA

AC	Y91243;
XX	
DT	22-MAY-2000 (first entry)
..	

Modified MvF Th epitope/LHRH antigenic peptide, SEQ ID NO:121.

KM Prominently T-cell epitope; measles virus F protein; WNV;
KM hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KM interleukin hormone releasing hormone; LHRH; contraceptive; anticancer
KM somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KM foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KM plasmidum falciparum; circumsporozoite; antimalarial; CEMP;
KM cholesterol ester transport protein; anti-arteriosclerotic.

05 Chimeric - Measles virus.

PN W09966957-A2

PD 29-DEC-1999

PF 21-JUN-1999; 99WO-US13975;

PR 20-JUN-1998; 98US-0100412.

PA (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY

DR WPI; 2000-160564/14.

PT New artificial T helper cell epitope and derived immunogens with target

PT or human immune deficiency virus -

PS Disclosure; Page 108; 129pp; English

CC The invention relates to novel promiscuous T helper cell epitopes (Th),

CC along with B cell epitopes. The Th epitopes and peptide immunogens containing them are used to induce Th1 and Th2 responses.

specifically against *Plasmodium falciparum*, cholesterol ester transport

CC Immunoreactive self-antigen or tumour antigen. The Th epitopes and

CC infections (HIV, foot-and-mouth disease or malaria); for cancer

immunotherapy: for inhibition of the action of luteinising hormone releasing hormone (LHRH) for contraception, treatment of hormone-dependent cancer, prevention of boar taint in meat, and Immunocastration): for promoting the growth of animals; or for treating allergies or arteriosclerosis. Incorporation of a promiscuous Th (functional in genetically diverse subjects) into an immunogen improves capacity to induce a strong T helper cell-mediated immune response, resulting in production of antibodies against a target antigen. Th can replace carrier proteins and pathogen-derived T helper epitopes. Sequence Y91121 represents a promiscuous T helper epitope from the measles virus F (MV) protein and sequences Y91122-Y91142, Y91226 and Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th epitope. Sequence Y91143 represents a promiscuous Th epitope from hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are synthetic epitopes derived from this HBV epitope. Y91156-Y91196, Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target antigenic peptide used in these LHRH antigenic peptides. Y91200 is somatostatin, and Y91201-Y91207 are antigenic peptides comprising somatostatin and a Th epitope. Somatostatin immunogens may be used to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain antigenic site, and Y91209-Y90211 are MVA Th epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV infection of T cells. Y90212 is a modified version of a human IGE (immunoglobulin E) CH3 domain, and Y90213-Y90219 are Th epitope/IGE CH3 antigenic peptides which may be used in the treatment of allergies. Y91220 is a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223 is a Plasmodium falciparum circumsporozoite (CS) target antigen, and Y91224-Y91225 comprise the CS antigen and an MVA Th epitope and may be used in a malaria vaccine. Y91228-Y91231 represent CERP-derived peptides and Y91232-Y91241 are immunogens comprising a CERP peptide and a Th epitope which may be used to prevent or treat arteriosclerosis and cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic peptides comprising MVA Th and HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively an immunostimulatory invasive protein epitope from *Yersinia* species, and hinge spacer peptide, both of which may optionally be used in the antigenic peptides of the invention.

Query Match	75.0%	Score 81	DB 21	Length 35
Best Local Similarity	81.2%	Pred. No. 0.00039		
Matches 13; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0

```
QY      2 PPPSSEHWSYGLRPG 17
          ||| | | | | | | |
Db     20 ppxpxpehwsyglrrpg 35
```

RESULT	8
Y68593	
ID	Y68593 standard; peptide; 35 AA.

AC	Y68593;
XX	
DT	05-MAY-2000 (first entry)

DE Peptide immunogen comprising a Th epitope and LHRH target antigen

KW Helper T cell epitope; F protein; Measles virus; peptide immunogen; LHRH;

oestrus; sexual development; sex hormone; promiscuous T helper epitope;

KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.

05 Synthetic.

OS Unidentified.

```

FH Key Location/Qualifiers
FT Peptide 1..19
FT /note= "helper Th epitope Y68549"
FT Misc-difference 4
FT /label= Ser, Thr
FT Misc-difference 7
FT /label= Lys, Arg
FT Misc-difference 8
FT /label= Gly, Thr
FT Misc-difference 12
FT /label= His, Thr
FT Misc-difference 13
FT /label= Lys, Arg
FT Misc-difference 16
FT /label= Gly, Thr
FT Peptide 20..25
FT /note= "spacer Y68564"
FT Misc-difference 22
FT /note= "not specified"
FT Misc-difference 24
FT /note= "not specified"
FT Peptide 26..35
FT /note= "LHRH antigenic epitope Y68566"
XX WO966952-A1.
XX 29-DEC-1999.
XX 21-JUN-1999; 99WO-US13960.
XX 20-JUN-1998; 98US-0100414.
XX (UNBI-) UNITED BIOMEDICAL INC.
XX Wang CY;
XX WPI; 2000-160562/14.
XX DR 2000-160562/14.
XX PT New peptide immunogen containing luteinising hormone-releasing hormone
XX PT antigen site and helper T cell epitope, for e.g. contraception and
XX PT treatment of cancer -
XX PS Example 1; Page 85; 102pp; English.
XX XX
XX CC The present sequence represents a peptide immunogen comprising a
XX CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
XX CC hormone-releasing hormone (LHRH). The synthetic Th epitope is derived
XX CC from a structured synthetic antigen library (SSAL) designated SSAL1 Th1.
XX CC SSAL Th1 is modelled after a promiscuous epitope taken from the F protein
XX CC of the Measles virus. The peptide immunogens cause induction of a
XX CC specific immune response to LHRH which is involved in regulation of a
XX CC spermatogenesis, ovulation, oestrus, sexual development and secretion
XX CC of sex hormones. Provision of a promiscuous T helper epitope (which is
XX CC functional in genetically diverse subjects) provides optimum
XX CC immunogenicity to the B cell epitopes of the target antigen and thus
XX CC high antibody titres against the target antigen. The peptide immunogens
XX CC of the invention are used to vaccinate against mammalian LHRH, for use
XX CC as (reversible) contraceptive; control of hormone-dependent tumours
XX CC (cancer of prostate or breast, also endometriosis); to prevent boar
XX CC taint (and improve meat quality) and for immunocastration.
XX XX
XX S0 Sequence 35 AA;

```

Query Match 75.0%; Score 81; DB 21; Length 35;
 Best Local Similarity 81.2%; Pred. No. 0.00039;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 2 PPPSSEHWSYGLRPG 17
  | | | | | | | | | |
Db 20 ppxpxpexwsyglrpg 35

```

```

RESULT 9
ID W03944
W03944 standard; Protein; 49 AA.
XX
XX AC W03944;
XX
XX D7 20-NOV-1996 (first entry)
XX
XX DE GNRH 4-repeat sequence.
XX
XX KW Leukotoxin; LKT; gonadotropin-releasing hormone; GNRH;
XX KW fusion protein; immunogen; vaccine; fertility control;
XX KW contraceptive; sterilisation; PCB113; PCB11.
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Peptide 1..10
XX FT /label= GNRH
XX FT Peptide 11..13
XX FT /label= Spacer
XX FT Peptide 14..23
XX FT /label= GNRH
XX FT Peptide 24..26
XX FT /label= Spacer
XX FT Peptide 27..36
XX FT /label= GNRH
XX FT Peptide 37..39
XX FT /label= Spacer
XX FT Peptide 40..49
XX FT /label= GNRH
XX
XX PN W09624675-A1.
XX
XX PD 15-AUG-1996.
XX
XX PE 24-JAN-1996; 96WO-CA00049.
XX
XX PR 10-FEB-1995; 95US-0387156.
XX
XX PA (UYSA-) UNIV SASKATCHEWAN.
XX
XX PI Manns JG, Potter AA;
XX
XX DR WPI; 1996-384447/38.
XX DR N-PSDB; T37178.
XX
XX PT Gonadotropin-releasing hormone multimer fusion proteins - with
XX PT leukotoxin polypeptide for increased immunogenicity, useful in
XX PT antifertility vaccine prodn.
XX
XX PS Example 2; Fig 1B; 87pp; English.
XX
XX CC A synthetic DNA sequence (T37178) codes for a gonadotropin
XX CC releasing hormone (GNRH) tetramer (W03944), in which the 4 GNRH
XX CC repeat units are separated by spacers designed to increase
XX CC immunogenicity. The DNA sequence was incorporated into vector
XX CC pAA352 (ATCC 68283), which contains a truncated leukotoxin
XX CC gene (LKT 352) derived from Pasteurella haemolytica, to
XX CC give plasmid PCB113 (T37176). Escherichia coli transformants
XX CC produce an LKT-GNRH fusion protein (see also W03942 and W03943)
XX CC useful as a vaccine for fertility control of domestic or farm animals.
XX
XX S0 Sequence 49 AA;

```

Query Match 65.7%; Score 71; DB 17; Length 49;
 Best Local Similarity 80.0%; Pred. No. 0.0099;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

QY 3 PPPSSEHWSYGLRPG 17
  | | | | | | | | | |
Db 22 ppgssghwsyglrpg 36

```



```

RESULT 10
W79567
ID W79567 standard; Protein: 49 AA.
XX
AC W79567;
XX
DT 24-DEC-1998 (first entry)
XX
DE GnRH-2.
XX
KM Gonadotropin releasing hormone; GnRH; decapeptide; hypothalamus; spacer;
KM LH; luteinizing hormone; FSH; follicle stimulating hormone; vertebrate;
KM pyroglu; chimera; leukotoxin polypeptide; multimer; vaccine; tumour;
KM immunogenic.
XX
OS Synthetic.
XX
PN MO9806848-A1.
XX
PD 19-FEB-1998.
XX
PF 08-AUG-1997; 97WO-CA00559.
XX
PR 09-AUG-1996; 96US-0694865.
XX
PA (UYSA-) UNIV SASKATCHEWAN.
XX
PI Manns JG, Potter AA;
XX
DR WPI: 1998-159540/14.
XX
N-PSDB: V61529.
XX
PT Chimeric protein of leukotoxin and gonadotropin releasing hormone -
PT useful for, e.g. preparation of vaccines for reduction of incidence
PT of mammary tumours in mammals
XX
PS Disclosure: Figure 1B; 118pp; English.
XX
XX The present sequence represents a recombinantly produced or chemically
CC synthesised gonadotropin releasing hormone-2 (GnRH-2) polypeptide, which
CC contains four copies of the GnRH decapeptide (V61528) and triplet amino
CC acid spacers between each of these sequences. This decapeptide is
CC secreted naturally by the hypothalamus which controls release of both the
CC luteinizing hormone (LH) and the follicle stimulating hormone (FSH) in
CC vertebrates. This sequence, as compared to the native peptide, has been
CC found to have an N-terminal Gln rather than a pyroglu residue, and also
CC contains substitutions at amino acid residues 15 and 41, whereby His is
CC replaced by Asp. This produces an alternating multimeric GnRH sequence
CC which is highly immunogenic that can be used in the construction of a
CC chimeric protein that comprises a leukotoxin polypeptide, several
CC multimers, and the GnRH sequence. The chimeric protein can be used as a
CC vaccine to help reduce the incidence of mammary tumours in a mammalian
CC individual.
XX
SQ Sequence 49 AA:

```

```

Query Match 65.7%; Score 71; DB 19; Length 49;
Best Local Similarity 80.0%; Pred. No. 0.0099;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 3 PPPSSRHWSYGLRPG 17
   | 11:|||||11111
DB 22 pygssqhwsgylrpg 36

```

```

RESULT 11
W61542
ID W61542 standard; Protein: 49 AA.
XX
AC W61542;

```

```

XX
DT 27-OCT-1998 (first entry)
XX
DE Peptide hormone GnRH-2 decapeptide (4 copies) fragment.
XX
KM GnRH; gonadotropin releasing hormone peptide hormone; leukotoxin;
KM immunisation; endogenous molecule; vaccine; ear; immunogen; carrier;
KM immune response; hormone receptor; cancerous cell; domestic animal;
KM porcine; bovine; luteinizing hormone; follicle stimulating hormone;
KM immunocastrate.
XX
OS Synthetic.
XX
PN WO9834639-A1.
XX
PD 13-AUG-1998.
XX
PF 04-FEB-1998; 98WO-CA00059.
XX
PR 05-FEB-1997; 97US-0036883.
XX
PA (BIOS-) BIOSTAR INC.
XX
PI Acres SD, Harland R, Manns JG;
XX
DR WPI: 1998-446952/38.
XX
N-PSDB: V45190.
XX
PT Immunisation against endogenous molecules by administering vaccine
PT to ear - useful to elicit efficient and uniform immune response
PT against e.g. gonadotropin releasing hormone to immunocastrate pigs
PT and cattle
XX
PS Example 1; Fig 1B; 61pp; English.
XX
XX This represents the amino acid sequence of the gonadotropin releasing
CC hormone (GnRH-2) decapeptide fragment used in the chimeric leukotoxin-
CC GnRH polypeptide gene fusions. This is used to exemplify the method of
CC invention of immunisation against endogenous molecules by administering
CC a vaccine which comprises an immunogen and a carrier to the ear of the
CC mammal. The method is useful for eliciting an efficient and uniform
CC immune response to block or suppress the activity of an endogenous
CC hormone, hormone receptor, agonist or antagonist in a vaccinated subject,
CC or to elicit an immune response against a targeted endogenous cell type
CC (e.g. a cancerous or otherwise diseased cell). It is especially useful
CC to reduce the levels of GnRH in domestic animals, especially in porcine
CC or bovine species. The use of GnRH immunogens in the vaccine reduces the
CC levels of luteinizing hormone and follicle stimulating hormone and helps
CC in immunocastrating the animal. Administration of vaccine compositions to
CC the ear instead of intramuscular administration into the neck increases
CC the efficiency of vaccination of mammals against endogenous immunogens,
CC and may increase uniformity of vaccine presentation since the ear is
CC relatively uniform from animal to animal.
XX
SQ Sequence 49 AA:

```

```

Query Match 65.7%; Score 71; DB 19; Length 49;
Best Local Similarity 80.0%; Pred. No. 0.0099;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 3 PPPSSRHWSYGLRPG 17
   | 11:|||||11111
DB 22 pygssqhwsgylrpg 36

```

```

RESULT 12
Y58363
ID Y58363 standard; Protein: 49 AA.
XX
AC Y58363;
XX
DT 27-MAR-2000 (first entry)

```

XX DE Four-copy gonadotropin-releasing hormone (GnRH) multimer.
 XX GNHR multimer: gonadotropin-releasing hormone; immunosterilisation;
 KM immunocotraception; vaccine; feline; canine; equine; cervine; ds
 XX OS Mammalia.
 OS Synthetic.
 XX MO9962545-A2.
 XX PN
 XX PD 09-DEC-1999.
 XX PE 28-MAY-1999; 99WO-CA00493.
 XX PR 04-JUN-1998; 98US-0088024.
 XX PR 06-MAY-1999; 99US-0306689.
 XX PA (BIOS-) BIOSTAR INC.
 XX PI Robbins SC;
 XX PI WPI: 2000-086857/07.
 XX DR N-PSDB: 255702.
 XX DR
 XX PT Hormone immunogens, analogues or antibodies used to manufacture
 PT vaccines for suppression of reproductive behavior and fertility in
 PT vertebrates -
 XX PS Claim 5; Fig 5B; 88pp; English.
 XX CC This sequence represents a four-copy gonadotropin-releasing
 CC hormone (GnRH) multimer, where the second and fourth GnRH sequence
 CC have a His to Asp substitution at position 2 of the GnRH sequence.
 CC The invention relates to GnRH immunogens, analogues or antibodies
 CC that cross-react with endogenous GnRH of a vertebrate. A specifically
 CC claimed immunogenic fusion protein (Y58361) comprises, in the N to
 CC C-terminal direction, a synthetic peptide sequence (Y58364), an eight
 CC copy GnRH multimer (composed of two copies of the 4xGnRH multimer
 CC sequence of Y58363), the LKT protein (which functions as a carrier
 CC protein), and a second eight copy GnRH multimer. The fusion protein may
 CC be used in a vaccine composition for prepubertal administration to a
 CC vertebrate subject to result in prolonged suppression of reproductive
 CC behaviour and/or fertility. GnRH immunogens, analogues or antibodies are
 CC used to manufacture a composition or vaccine for immunosterilisation or
 CC immunocotraception of feline, canine, equine or cervine subjects.
 CC The vaccines are used to suppress reproductive behaviour and/or
 CC fertility for at least 10 months. The prepubertal administration
 CC results in a prolonged, long-term suppression of testicular development
 CC and/or function in males, or a prolonged, long-term suppression of
 CC ovarian development and/or function in females. The methods provide a
 CC viable and desirable alternative to surgical forms of sterilisation that
 CC are currently used.
 XX SQ Sequence 49 AA;
 OY Query Match 65.7%; Score 71; DB 21; Length 49;
 Best Local Similarity 80.0%; Pred. No. 0.0099;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 3 PPSSEHWSYGLRPG 17
 Db 22 pgssgqhsyglrpg 36
 RESULT 13
 Y58135
 XX ID Y58135 standard; Protein; 49 AA.
 XX AC Y58135;
 XX DT 07-MAR-2000 (first entry)
 DT

XX DE GnRH analogue multimer, containing four copies of the GnRH analogue.
 XX GNHR releasing hormone; GnRH; leukotoxin; LKT; fusion protein;
 KM antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;
 KM non-androgenic; steroid; reduction; weight gain; muscle distribution;
 KM fat distribution; male pattern; boar taint; flavour; impairment;
 KM reliable; immunocastration; meat production.
 XX OS Synthetic.
 OS Mammalia.
 XX MO9956771-A2.
 XX PN
 XX PD 11-NOV-1999.
 XX PE 05-MAY-1999; 99WO-CA00360.
 XX PR 05-MAY-1998; 98US-0084217.
 XX PA (BIOS-) BIOSTAR INC.
 XX PI Manns JG, Acres SD, Harland R;
 XX PI WPI: 2000-062125/05.
 XX DR N-PSDB: 246402.
 XX DR
 XX PT Production of uncastrated male food animals using vaccines -
 PT Example 1; Fig 2B; 87pp; English.
 XX PS
 XX CC This sequence represents four copies of a gonadotropin
 CC releasing hormone (GnRH) analogue, DNA encoding which was
 CC used in the construction of a chimeric GnRH-leukotoxin (LKT)
 CC fusion gene (Z46400). This fusion gene encodes a GnRH-LKT fusion
 CC protein which may be used as a vaccine. The LKT portion of the protein
 CC acts to enhance the immunogenicity of the GnRH portion. The invention
 CC relates to a method of using two GnRH immunogen vaccines to produce
 CC uncastrated male animals for meat production, one vaccination prior to
 CC or during the fattening period to reduce circulating testosterone levels,
 CC and the second vaccination about 2-8 weeks before slaughter to
 CC substantially reduce androgenic and/or non-androgenic steroids. The
 CC invention is used to produce food animals that exhibit the weight gain
 CC and muscle/fat distribution of male animals without the problems
 CC associated with male animals. Such problems include "boar taint", a
 CC urine-like odour found in cooked meat of uncastrated pigs which is
 CC caused by steroids stored in the tissues, and similar flavour
 CC impairments in the meat of other intact male animals. The invention is
 CC more reliable than prior art immunocastration techniques.
 XX SQ Sequence 49 AA;
 OY Query Match 65.7%; Score 71; DB 21; Length 49;
 Best Local Similarity 80.0%; Pred. No. 0.0099;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 3 PPSSEHWSYGLRPG 17
 Db 22 pgssgqhsyglrpg 36
 RESULT 14
 W03943
 XX ID W03943 standard; Protein; 544 AA.
 XX AC W03943;
 XX DT 20-NOV-1996 (first entry)
 XX DE LKT-GnRH protein fusion from pCB11.
 XX DT Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
 KM

KW		fusion protein; immunogen; vaccine; fertility control;
KW		contraceptive; sterilisation.
XX		
OS		Chimeric Pasteurella haemolytica A1 strain B122;
OS		Chimeric synthetic.
XX		
FH	Key	Location/Qualifiers
FT	Domain	1..493
FT	/Label= LKT	
FT	Domain	494..544
FT	/Label= GnRH_repeat_domain	
PX		
PN	M09624675-A1.	
XX		
PD	15-AUG-1996.	
XX		
PF	24-JAN-1996;	96WO-CA00049.
XX		
PR	10-FEB-1995;	95US-0387156.
XX		
PA	(UYSA-) UNIV SASKATCHEWAN.	
XX		
PI	Manns JG, Potter AA;	
DR	WPt: 1996-384447/38.	
N-PSDB:	T37177.	
XX		
PT	Gonadotropin-releasing hormone multimer fusion proteins - with	
PT	leukotoxin polypeptide for increased immunogenicity, useful in	
PT	antifertility vaccine prodn.	
XX		
PS	Claim 8; Fig 7A-7E; 87pp; English.	
XX		
CC	A chimeric protein (M03943) is composed of a fusion between	
CC	a truncated leukotoxin (LKT-111) from Pasteurella haemolytica	
CC	and a 4-copy gonadoliberin-releasing hormone (GnRH) repeat	
CC	sequence (see also M03944). It is the product of a chimeric	
CC	gene (T37177) produced by deleting an approx. 1300 bp sequence	
CC	from PCB113 (see also T37176) coding for amino acids 352-784	
CC	of LKT-352. Recombinant plasmid PCB11 (LKT 111:4 copy GnRH,	
CC	ATCC 69748) was obtd. Escherichia coli transformants produced	
CC	the chimeric protein, which is useful as a vaccine for fertility	
CC	control, esp. immunological sterilisation of domestic or farm	
CC	animals.	
SQ	Sequence	544 AA;
XX		
XX		
Query Match	65.7%;	Score 71; DB 17; Length 544;
Best Local Similarity	80.0%;	Pred. No. 0.099;
Matches 12; Conservative	1;	Mismatches 2; Indels 0; Gaps 0;
OY	3 PPPSEHHVSYGRPG 17	
	:	
Db	515 pgyssqhwsgylrpg 529	
RESULT 15		
W79570		
ID W79570 standard; Protein: 544 AA.		
XX		
AC W79570;		
XX		
DT 24-DEC-1998 (first entry)		
XX		
LKT-GnRH chimeric protein.		
XX		
Chimera; PCB11; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;		
KW cytotoxic activity; antigen presentation; immune response; vaccine;		
tumour.		
XX		
Synthetic.		
XX		

Query Match	Best Local Similarity	Score 71;	DB 19;	Length 544;
Matches 12;	Conservative 1;	Mismatches 2;	Indels 0;	Gaps 0
07 . 3 PPPSSHHWYGLRPG 17	1 11:111111111			
DB 515 pygssqhwswyglrpt 529				

Search completed: March 2, 2001, 10:53:46
Job time: 911 sec

	RESULT	15
W79570		
ID	W79570 standard; Protein; 544 AA.	
XX		
XX		
AC	W79570;	
XX		
XX		
DT	24-DEC-1998 (first entry)	
XX		
XX		
DE	LKT-GnRH chimeric protein.	
XX		
KW	Chimera; pCB11; LKT 11; GnRH; Gonadotropin releasing hormone; multimer.	
KW	Cytotoxic activity; antigen presentation; immune response; vaccine;	
KW	tumour.	
XX		
XX		
OS	Synthetic.	
XX		

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:59:02 ; Search time 201.99 Seconds

(without alignments)
12.920 Million cell updates/sec

Title: US-09-306-689-3

Perfect score: 108

Sequence: 1 CPPPSEHWSYGLRPG 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 957798 seqs, 153513742 residues

Total number of hits satisfying chosen parameters: 957798

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA:*

1: /cgn2_6/ptoddata/2/paa/PCURS.COMB.pep.*
2: /cgn2_6/ptoddata/2/paa/US06.COMB.pep.*
3: /cgn2_6/ptoddata/2/paa/US07.COMB.pep.*
4: /cgn2_6/ptoddata/2/paa/US08.COMB.pep.*
5: /cgn2_6/ptoddata/2/paa/US081.COMB.pep.*
6: /cgn2_6/ptoddata/2/paa/US082.COMB.pep.*
7: /cgn2_6/ptoddata/2/paa/US083.COMB.pep.*
8: /cgn2_6/ptoddata/2/paa/US084.COMB.pep.*
9: /cgn2_6/ptoddata/2/paa/US085.COMB.pep.*
10: /cgn2_6/ptoddata/2/paa/US086.COMB.pep.*
11: /cgn2_6/ptoddata/2/paa/US087.COMB.pep.*
12: /cgn2_6/ptoddata/2/paa/US088.COMB.pep.*
13: /cgn2_6/ptoddata/2/paa/US089.COMB.pep.*
14: /cgn2_6/ptoddata/2/paa/US090.COMB.pep.*
15: /cgn2_6/ptoddata/2/paa/US091.COMB.pep.*
16: /cgn2_6/ptoddata/2/paa/US092.COMB.pep.*
17: /cgn2_6/ptoddata/2/paa/US093.COMB.pep.*
18: /cgn2_6/ptoddata/2/paa/US094.COMB.pep.*
19: /cgn2_6/ptoddata/2/paa/US095.COMB.pep.*
20: /cgn2_6/ptoddata/2/paa/US096.COMB.pep.*
21: /cgn2_6/ptoddata/2/paa/US097.COMB.pep.*
22: /cgn2_6/ptoddata/2/paa/US098.COMB.pep.*
23: /cgn2_6/ptoddata/2/paa/US099.COMB.pep.*
24: /cgn2_6/ptoddata/2/paa/US100.COMB.pep.*
25: /cgn2_6/ptoddata/2/paa/US101.COMB.pep.*
26: /cgn2_6/ptoddata/2/paa/US102.COMB.pep.*
27: /cgn2_6/ptoddata/2/paa/US103.COMB.pep.*
28: /cgn2_6/ptoddata/2/paa/US104.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	100.0	17	US-09-305-924-3	Sequence 3, Appli
2	108	100.0	17	US-09-306-689-3	Sequence 3, Appli
3	102	94.4	17	PCT-US95-01225-5	Sequence 5, Appli
4	102	94.4	17	US-08-478-546B-5	Sequence 5, Appli
5	82.5	76.4	16	PCT-US95-01225-7	Sequence 7, Appli

6	82.5	76.4	16	8	US-08-478-546B-7	Sequence 7, Appli
7	81	75.0	34	1	PCT-US99-13960-56	Sequence 56, Appli
8	81	75.0	34	1	PCT-US99-13975B-122	Sequence 122, Appli
9	81	75.0	34	15	US-09-100-414-56	Sequence 56, Appli
10	81	75.0	34	15	US-09-100-414A-82	Sequence 82, Appli
11	81	75.0	34	17	US-09-303-323-82	Sequence 122, Appli
12	81	75.0	34	27	US-09-701-588-122	Sequence 55, Appli
13	81	75.0	35	1	PCT-US99-13960-55	Sequence 120, Appli
14	81	75.0	35	1	PCT-US99-13975B-120	Sequence 121, Appli
15	81	75.0	35	15	US-09-100-414-55	Sequence 55, Appli
16	81	75.0	35	15	US-09-100-414A-80	Sequence 80, Appli
17	81	75.0	35	15	US-09-100-414A-81	Sequence 81, Appli
18	81	75.0	35	17	US-09-303-323-80	Sequence 81, Appli
19	81	75.0	35	17	US-09-303-323-81	Sequence 121, Appli
20	81	75.0	35	17	US-09-303-323-81	Sequence 121, Appli
21	81	75.0	35	27	US-09-701-588-120	Sequence 121, Appli
22	81	75.0	35	27	US-09-701-588-121	Sequence 121, Appli
23	72.5	67.1	397	19	US-09-506-078-31	Sequence 31, Appli
24	72.5	67.1	397	22	US-60-120-454-31	Sequence 25, Appli
25	72.5	67.1	398	22	US-60-120-454-25	Sequence 25, Appli
26	72.5	67.1	398	22	US-60-120-454-25	Sequence 25, Appli
27	72.5	67.1	442	19	US-09-506-078-27	Sequence 27, Appli
28	72.5	67.1	442	22	US-60-120-454-27	Sequence 27, Appli
29	71	65.7	49	14	US-09-019-010-4	Sequence 4, Appli
30	71	65.7	49	16	US-09-249-447A-7	Sequence 7, Appli
31	71	65.7	49	17	US-09-305-924-11	Sequence 11, Appli
32	71	65.7	49	17	US-09-306-689-11	Sequence 11, Appli
33	71	65.7	49	17	US-09-383-912-4	Sequence 4, Appli
34	71	65.7	544	17	US-09-383-912-10	Sequence 10, Appli
35	71	65.7	695	17	US-09-305-924-13	Sequence 13, Appli
36	71	65.7	695	17	US-09-306-689-13	Sequence 13, Appli
37	71	65.7	699	17	US-09-383-912-16	Sequence 16, Appli
38	71	65.7	977	17	US-09-383-912-8	Sequence 8, Appli
39	69	63.9	49	22	US-60-192-739-3463	Sequence 3463, Ap
40	68	63.0	14	5	US-08-160-882-24	Sequence 24, Appli
41	68	63.0	24	5	US-08-160-882-43	Sequence 43, Appli
42	68	63.0	44	5	US-08-160-882-45	Sequence 45, Appli
43	68	63.0	84	5	US-08-160-882-47	Sequence 47, Appli
44	68	63.0	90	13	US-08-912-314A-19	Sequence 19, Appli
45	66	61.1	17	5	US-08-160-882-16	Sequence 16, Appli

ALIGNMENTS

RESULT 1
US-09-305-924-3
; Sequence 3, Application US/09305924A
; GENERAL INFORMATION:
; APPLICANT: Jack G. Manns
; APPLICANT: Stephen D. Acres
; APPLICANT: Richard Hatland
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305, 924A
; EARLIER FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084, 217
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 17
; TYPE: PRT
; ORGANISM: GNRH
US-09-305-924-3

Query Match 100.0% Score 108: DB 17: Length 17: 1
Best Local Similarity 100.0% Pred. No. 8.2e-07: 1
Matches 17: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
1 CPPPSEHWSYGLRPG 17
|||||

Db 1 CPPPSEHWSYGLRPG 17

RESULT 2

US-09-306-689-3

Sequence 3, Application US/09306689B

GENERAL INFORMATION:

APPLICANT: Robbins, Sarah C.

TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN

TITLE OF INVENTION: ANIMALS

FILE REFERENCE: 9001-0047

CURRENT APPLICATION NUMBER: US/09/306,689B

CURRENT FILING DATE: 1999-05-06

EARLIER APPLICATION NUMBER: US 60/088,024

EARLIER FILING DATE: 1998-06-04

NUMBER OF SEQ ID NOS: 14

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 17

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Sequence

US-09-306-689-3

Query Match 100.0%; Score 108; DB 17; Length 17;

Best Local Similarity 100.0%; Pred. No. 8,2e-07;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 CPPPSEHWSYGLRPG 17

Db 1 CPPPSEHWSYGLRPG 17

RESULT 3

PCT-US95-01225-5

Sequence 5, Application PC/RUS9501225

GENERAL INFORMATION:

APPLICANT: Grimes, Stephen

APPLICANT: Scibienski, Robert

TITLE OF INVENTION: Immunogens Against Gonadotropin

TITLE OF INVENTION: Releasing Hormone

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dimitrios T. Drivas, Esq.

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036-2787

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/01225

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Drivas Esq., Dimitrios T.

REGISTRATION NUMBER: 32,218

REFERENCE/DOCKET NUMBER: 1102865-300

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-819-8286

TELEFAX: 212-354-8113

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: YES

FEATURE:

NAME/KEY: Region

LOCATION: 1..17

OTHER INFORMATION: /note= "spacer"

FEATURE:

NAME/KEY: Region

LOCATION: 8..17

OTHER INFORMATION: /note= "immunomimic"

FEATURE:

NAME/KEY: Modified-site

LOCATION: 17

OTHER INFORMATION: /label= glyNH2

OTHER INFORMATION: /note= "glycinamide"

PCT-US95-01225-5

Query Match 94.4%; Score 102; DB 1; Length 17;

Best Local Similarity 100.0%; Pred. No. 4,5e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 CPPPSEHWSYGLRP 16

Db 1 CPPPSEHWSYGLRP 16

RESULT 4

US-08-478-546B-5

Sequence 5, Application US/08478546B

GENERAL INFORMATION:

APPLICANT: Grimes, Stephen

APPLICANT: Scibienski, Robert

TITLE OF INVENTION: Methods for the Treatment of Hormone-Dependent

TITLE OF INVENTION: Tumors with Immunogens against Gonadotropin Releasing Hormo

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dimitrios T. Drivas, Esq.

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036-2787

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/478,546B

FILING DATE: 07-JUN-1995

CLASSIFICATION: 424

Prior APPLICATION DATA:

APPLICATION NUMBER: US 08/188,223

FILING DATE: 27-JAN-1994

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Drivas Esq., Dimitrios T.

REGISTRATION NUMBER: 32,218

REFERENCE/DOCKET NUMBER: 1102865-300

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-819-8286

TELEFAX: 212-354-8113

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Region

LOCATION: 1..17

OTHER INFORMATION: /note= "spacer"

FEATURE:
NAME/KEY: Region
LOCATION: 8..17
OTHER INFORMATION: /note="immunomimic"
FEATURE:
NAME/KEY: Modified site
LOCATION: 17
OTHER INFORMATION: /note="Xaa"
OTHER INFORMATION: /note="amidated glycine"
US-08-478-546B-5

Query Match 94.4%; Score 102; DB 8; Length 17;
Best Local Similarity 100.0%; Pred. No. 4.5e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CPPPSSEHWSYGLRP 16
DB 1 CPPPSSEHWSYGLRP 16

RESULT 5
PCT-US95-01225-7
Sequence 7, Application PC/TUS9501225
GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
TITLE OF INVENTION: Immunogens Against Gonadotropin
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dimitrios T. Drivas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01225
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8286
TELEFAX: 212-354-8113
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
FEATURE:
NAME/KEY: Region
LOCATION: 1..6
OTHER INFORMATION: /note="spacer"
FEATURE:
NAME/KEY: Region
LOCATION: 7..16
OTHER INFORMATION: /note="immunomimic"
FEATURE:
NAME/KEY: Modified site
LOCATION: 16
OTHER INFORMATION: /label="GlyNH2"
OTHER INFORMATION: /note="glycinamide"

PCT-US95-01225-7

Query Match 76.4%; Score 82.5; DB 1; Length 16;
Best Local Similarity 87.5%; Pred. No. 0.0011;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CPPPSSEHWSYGLRP 16
DB 1 CPPPP-REHWSYGLRP 15

RESULT 6
US-08-478-546B-7
Sequence 7, Application US/08478546B
GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
TITLE OF INVENTION: Methods for the Treatment of Hormone-Dependent
Tumors with Immunogens against Gonadotropin Releasing Hormo
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dimitrios T. Drivas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,546B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/188,223
FILING DATE: 27-JAN-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8286
TELEFAX: 212-354-8113
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Region
LOCATION: 1..6
OTHER INFORMATION: /note="spacer"
FEATURE:
NAME/KEY: Region
LOCATION: 7..16
OTHER INFORMATION: /note="immunomimic"
FEATURE:
NAME/KEY: Modified site
LOCATION: 16
OTHER INFORMATION: /note="Xaa"
OTHER INFORMATION: /note="amidated glycine"
US-08-478-546B-7

Query Match 76.4%; Score 82.5; DB 8; Length 16;
Best Local Similarity 87.5%; Pred. No. 0.0011;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 1 CPPPSEHWSYGLRP 16
| | | | |
Db 1 CPPP-REHWSYGLRP 15

RESULT 7

PCT-US99-13960-56
; Sequence 56, Application PC/TUS9913960

GENERAL INFORMATION:

APPLICANT: UNITED BIOMEDICAL, INC.

TITLE OF INVENTION: ARTIFICIAL T HELPER CELL EPITOPES

TITLE OF INVENTION: AS IMMUNE STIMULATORS FOR SYNTHETIC PEPTIDE IMMUNOGENS INCLUDED

TITLE OF INVENTION: IMMUNOGENIC LHRH PEPTIDES

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan & Finnegan, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0054

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC Windows

SOFTWARE: Word 97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US99/13960

FILING DATE: 21-JUNE-1999

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/100,414

FILING DATE: 20-JUNE-1998

ATTORNEY/AGENT INFORMATION:

NAME: Maria C.H. Lin

REGISTRATION NUMBER: 29,323

REFERENCE/DOCKET NUMBER: 1151-4157PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-758-4800

TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO: 56:

SEQUENCE CHARACTERISTICS:

LENGTH: 34 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

PCT-US99-13960-56

Query Match 75.0%; Score 81; DB 1; Length 34;

Best Local Similarity 81.2%; Pred. No. 0.0033;

Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 PPPPSEHWSYGLRP 17
| | | | |
Db 19 PPXPPEHWSYGLRP 34

RESULT 8

PCT-US99-13975B-122

; Sequence 122, Application PC/TUS9913975B

GENERAL INFORMATION:

APPLICANT: UNITED BIOMEDICAL INC., ET AL.

TITLE OF INVENTION: ARTIFICIAL T HELPER CELL

TITLE OF INVENTION: EPITOPES AS IMMUNE STIMULATORS FOR SYNTHETIC

PEPTIDE IMMUNOGENS

NUMBER OF SEQUENCES: 151

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan & Finnegan, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0054

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC Windows

SOFTWARE: Word 97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US99/13975B

FILING DATE: 21-JUNE-1999

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/100,412

FILING DATE: 20-JUNE-1998

ATTORNEY/AGENT INFORMATION:

NAME: Maria H. Lin

REGISTRATION NUMBER: 29,323

REFERENCE/DOCKET NUMBER: 1151-4158PC1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-758-4800

TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO: 122:

SEQUENCE CHARACTERISTICS:

LENGTH: 34 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

PCT-US99-13975B-122

Query Match 75.0%; Score 81; DB 1; Length 34;
Best Local Similarity 81.2%; Pred. No. 0.0033;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 PPPPSEHWSYGLRP 17
| | | | |
Db 19 PPXPPEHWSYGLRP 34

RESULT 9

US-09-100-414-56

; Sequence 56, Application US/09100414

GENERAL INFORMATION:

APPLICANT: Wang, Chang Y1

TITLE OF INVENTION: NOVEL LHRH PEPTIDE

TITLE OF INVENTION: IMMUNOGENS

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan & Finnegan, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0054

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC Windows

SOFTWARE: Word 97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/100,414

FILING DATE: 20-JUNE-1998

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Maria H. Lin

REGISTRATION NUMBER: 29,323

REFERENCE/DOCKET NUMBER: 1151-4157

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-758-4800

TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO: 56:

SEQUENCE CHARACTERISTICS:

LENGTH: 34 amino acids

TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414-56

Query Match 75.0%; Score 81; DB 15; Length 34;
Best Local Similarity 81.2%; Pred. No. 0.0033;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PPPSSEHWSYGLRPG 17
Db 19 PPXPPEHWSYGLRPG 34

RESULT 10
US-09-100-414A-82
Sequence 82, Application US/09100414A
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414A
FILING DATE: 20-JUNE-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414A-82

Query Match 75.0%; Score 81; DB 15; Length 34;
Best Local Similarity 81.2%; Pred. No. 0.0033;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PPPSSEHWSYGLRPG 17
Db 19 PPXPPEHWSYGLRPG 34

RESULT 11
US-09-303-323-82
Sequence 82, Application US/09303323
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.

STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-82

Query Match 75.0%; Score 81; DB 17; Length 34;
Best Local Similarity 81.2%; Pred. No. 0.0033;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PPPSSEHWSYGLRPG 17
Db 19 PPXPPEHWSYGLRPG 34

RESULT 12
US-09-701-588-122
Sequence 122, Application US/09701588
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF SEQUENCES: 122
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/701,588
FILING DATE: 29-NOV-2000
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-158051

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 122:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 122:
US-09-701-588-122

Query Match 75.0%; Score 81; DB 27; Length 34;
Best Local Similarity 81.2%; Pred. No. 0.0033;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 PPPSSEHWSYGLRPG 17
||| |||||
Db 19 PPXPPEHWSYGLRPG 34

RESULT 13
PCT-US99-13960-55
Sequence 55, Application PC/TUS9913960
GENERAL INFORMATION:
APPLICANT: UNITED BIOMEDICAL INC.
TITLE OF INVENTION: ARTIFICIAL T HELPER CELL EPITOPES
TITLE OF INVENTION: AS IMMUNE STIMULATORS FOR SYNTHETIC PEPTIDE IMMUNOGENS INCLUDED
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US99/13960
FILING DATE: 21-JUNE-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria C.H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 4
OTHER INFORMATION: /note= "Ser or
OTHER INFORMATION: Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 7
OTHER INFORMATION: /note= "Lys or
OTHER INFORMATION: Arg"
FEATURE:

NAME/KEY: Modified-site
LOCATION: 8
OTHER INFORMATION: /note= "Gly or
OTHER INFORMATION: Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 12
OTHER INFORMATION: /note= "His or
OTHER INFORMATION: Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 13
OTHER INFORMATION: /note= "Lys or
OTHER INFORMATION: Arg"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 16
OTHER INFORMATION: /note= "Gly or
OTHER INFORMATION: Thr"
PCT-US99-13960-55

Query Match 75.0%; Score 81; DB 1; Length 35;
Best Local Similarity 81.2%; Pred. No. 0.0034;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 PPPSSEHWSYGLRPG 17
||| |||||
Db 20 PPXPPEHWSYGLRPG 35

RESULT 14
PCT-US99-13975B-120
Sequence 120, Application PC/TUS9913975B
GENERAL INFORMATION:
APPLICANT: UNITED BIOMEDICAL INC., ET AL.
TITLE OF INVENTION: ARTIFICIAL T HELPER CELL
TITLE OF INVENTION: EPITOPES AS IMMUNE STIMULATORS FOR SYNTHETIC
NUMBER OF SEQUENCES: 151
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US99/13975B
FILING DATE: 21-JUNE-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,412
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4158PC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 120:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US99-13975B-120

Query Match 75.0%; Score 81; DB 1; Length 35;
 Best Local Similarity 81.2%; Pred. NO. 0.0034;
 Matches 13; Conservative 0; Mismatches 3; Indels 0;

OY 2 PPPSSEHWSYGLRPG 17
 ||| |||||
 Db 20 PPXPPEHWSYGLRPG 35

RESULT 15

PCT-US99-13975B-121
 ; Sequence 121, Application PC/TUS9913975B
 ; GENERAL INFORMATION:
 ; APPLICANT: UNITED BIOMEDICAL INC., ET AL.
 ; TITLE OF INVENTION: ARTIFICIAL T HELPER CELL
 ; TITLE OF INVENTION: EPITOPES AS IMMUNE STIMULATORS FOR SYNTHETIC
 ; TITLE OF INVENTION: PEPTIDE IMMUNOGENS
 ; NUMBER OF SEQUENCES: 151
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Morgan & Finnegan, L.L.P.
 ; STREET: 345 Park Avenue
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10154-0054
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC Windows
 ; SOFTWARE: Word 97
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US99/13975B
 ; FILING DATE: 21-JUNE-1999
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/100,412
 ; FILING DATE: 20-JUNE-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Maria H. Lin
 ; REGISTRATION NUMBER: 29,323
 ; REFERENCE/DOCKET NUMBER: 1151-4158PCL
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-758-4800
 ; TELEFAX: 212-751-6849
 ; INFORMATION FOR SEQ ID NO: 121:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 35 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; PCT-US99-13975B-121

Query Match 75.0%; Score 81; DB 1; Length 35;
 Best Local Similarity 81.2%; Pred. NO. 0.0034;
 Matches 13; Conservative 0; Mismatches 3; Indels 0;

OY 2 PPPSSEHWSYGLRPG 17
 ||| |||||
 Db 20 PPXPPEHWSYGLRPG 35

Search completed: March 2, 2001, 10:59:03
 Job time: 373 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:53:46 ; Search time 66.16 Seconds
(without alignments)
8.786 Million cell updates/sec

Title: US-09-306-689-4

Perfect score: 104

Sequence: 1 XHMSYGLRPGSSPPPC 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_36:*

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103	99.0	17	16 R78283	GnRH immunomimic a
2	103	99.0	17	21 Y58140	Gonadotropin relea
3	83.5	80.3	16	16 R78285	GnRH immunomimic a
4	83.5	80.3	16	21 Y58141	Gonadotropin relea
5	59.6	59.6	14	15 R44953	E. coli fimbriae H
6	59.6	59.6	49	17 W03944	GnRH 4-repeat sequ
7	59.6	59.6	49	19 W79567	GnRH-2. Synthetic
8	59.6	59.6	49	19 W61542	Peptide hormone Gn
9	59.6	59.6	49	21 Y58363	Four-copy gonadotr
10	59.6	59.6	49	21 Y58135	GnRH analogue mult
11	62	59.6	257	12 R11178	Plasmid pBTA735-en
12	62	59.6	257	12 R11183	Plasmid pBTA740-en

13	62	59.6	544	17 W03943	LKT-GnRH protein f
14	62	59.6	544	19 W79570	LKT-GnRH chimeric
15	62	59.6	695	19 W79573	LKT-GnRH chimeric
16	62	59.6	695	21 Y58361	Leukotoxin/gonadot
17	62	59.6	695	21 Y58133	Gonadotropin relea
18	62	59.6	977	17 W03942	LKT-GnRH protein f
19	62	59.6	977	17 W79569	LKT-GnRH chimeric
20	59	56.7	12	15 R44954	E. coli fimbriae H
21	59	56.7	18	21 Y89763	Core polypeptide f
22	59	56.7	253	12 R11180	Plasmid pBTA734-en
23	59	56.7	253	12 R11181	Plasmid pBTA733-en
24	59	56.7	253	12 R11182	Plasmid pBTA731-en
25	59	56.7	263	12 R11185	Plasmid pBTA70-en
26	59	56.7	283	12 R11186	Plasmid pBTA62-en
27	59	56.7	323	12 R11187	Plasmid pBTA59-en
28	58	55.8	9	20 W94891	LHRH peptide fragm
29	58	55.8	10	2 P10097	Sequence of lutein
30	58	55.8	10	2 P10411	Luteinizing Hormon
31	58	55.8	10	2 P10416	Luteinizing Hormon
32	58	55.8	10	6 P50222	Gonadotropin rele
33	58	55.8	10	7 P60127	Gonadotropin relea
34	58	55.8	10	7 P61403	Gonadotropin relea
35	58	55.8	10	7 P60576	Novel decapeptide
36	58	55.8	10	8 P70922	Luteinizing hormon
37	58	55.8	10	10 P90630	Sequence of lutein
38	58	55.8	10	12 R15713	Peptide #1 with ho
39	58	55.8	10	13 R26819	LH releasing hormo
40	58	55.8	10	15 R62689	LHRH hapten for at
41	58	55.8	10	16 R91197	LHRH peptide. Syn
42	58	55.8	10	16 R86845	Gonadotropin relea
43	58	55.8	10	16 R75152	Gonadotropin relea
44	58	55.8	10	17 W65201	Luteinizing hormon
45	58	55.8	10	17 Y55061	Luteinizing hormon

ALIGNMENTS

RESULT 1	
ID R78283	standard; peptide; 17 AA.
XX	
AC R78283;	
XX	
DT 13-MAR-1996	(first entry)
XX	
DE GnRH immunomimic and spacer (GnRH(1-10)-Ser10).	
KW	Immunomimetic; gonadotropin releasing hormone; GnRH; LHRH; vaccine;
KW	luteinizing hormone releasing hormone; spacer; immunomimic; uterine;
KW	diphtheria toxin; DT; gynaecological; endometriosis; uterine fibroids;
KW	gonadal steroid hormone associated dependent disease; gonadotropin;
KW	immunological contraception; mammal; breast; cancer; prostate;
KW	benign prostatic hypertrophy.
XX	
OS Chimeric - Homo sapiens.	
OS	
XX	
OS Chimeric - Synthetic.	
XX	
FH Key	Location/Qualifiers
FT Modified-site	1
FT	/label= OTHER
FT	/note= "pyroglutamic acid"
FT	Peptide
FT	1..10
FT	/note= "GnRH"
FT	11..17
FT	/note= "spacer"
XX	
PN W09520600-A1.	
XX	
PD 03-AUG-1995.	
XX	
PF 26-JAN-1995;	95WO-US01225.
XX	

CC The sequence is that of an insert contg. the decapeptide gonadotropin
CC releasing hormone (GnRH) located in the major subunit of P-fimbriae
CC with serotype III at a position corresponding to a position in
CC hypervariable region 4 of the wild type major subunit. This forms
CC part of an immunogenic carrier system capable of eliciting an immune
CC response against GnRH.

XX Sequence 14 AA;

Query Match 59.6%; Score 62; DB 15; Length 14;

Best Local Similarity 100.0%; Pred. No. 0.066;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGS 11
| | | | | | | | | |

Db 3 hwsyglrpgs 12

RESULT 6

W03944 W03944 standard; Protein; 49 AA.

XX W03944;

XX 20-NOV-1996 (first entry)

XX GnRH 4-repeat sequence.

XX Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;

XX fusion protein; immunogen; vaccine; fertility control;

XX contriceptive; sterilisation; PCB113; PCB111.

XX Synthetic.

XX Key Location/Qualifiers

FT Peptide 1..10

FT Peptide /label- GnRH

FT Peptide 11..13

FT Peptide /label- Spacer

FT Peptide 14..23

FT Peptide /label- GnRH

FT Peptide 24..26

FT Peptide /label- Spacer

FT Peptide /label- GnRH

FT Peptide 37..39

FT Peptide /label- Spacer

FT Peptide 40..49

FT Peptide /label- GnRH

XX W09624675-A1.

XX 15-AUG-1996.

XX 24-JAN-1996; 96WO-CA00049.

XX 10-FEB-1995; 95US-0387156.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Manns JG, Potter AA;

XX WPI: 1996-384447/38.

XX N-PSDB: T37178.

PT Gonadotropin-releasing hormone multimer fusion proteins - with
PT leukotoxin polypeptide for increased immunogenicity, useful in
PT antifertility vaccine prodn.

XX Example 2; Fig 1B; 87pp; English.

XX A synthetic DNA sequence (T37178) codes for a gonadotropin

CC releasing hormone (GnRH) tetramer (W03944), in which the 4 GnRH
CC repeat units are separated by spacers designed to increase
CC immunogenicity. The DNA sequence was incorporated into vector
CC pAA352 (ATCC 68283), which contains a truncated leukotoxin
CC gene (LKT 352) derived from Pasteurella haemolytica, to
CC give plasmid PCB113 (T37176). Escherichia coli transformants
CC produce an LKT-GnRH fusion protein (see also W03942 and W03943)
CC useful as a vaccine for fertility control of domestic or farm animals.

XX Sequence 49 AA;

Query Match 59.6%; Score 62; DB 17; Length 49;

Best Local Similarity 100.0%; Pred. No. 0.21;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGS 11
| | | | | | | | | |

Db 2 hwsyglrpgs 11

RESULT 7

W79567 W79567 standard; Protein; 49 AA.

XX W79567;

XX 24-DEC-1998 (first entry)

XX GnRH-2.

XX Gonadotropin releasing hormone; GnRH; decapeptide; hypothalamus; spacer;

XX LH; luteinising hormone; FSH; follicle stimulating hormone; vertebrate;

XX pyrogliu; chimera; leukotoxin polypeptide; multimer; vaccine; tumour;

XX immunogenic.

XX Synthetic.

XX W09806848-A1.

XX 19-FEB-1998.

XX 08-AUG-1997; 97WO-CA00559.

XX 09-AUG-1996; 96US-0694865.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Manns JG, Potter AA;

XX WPI: 1998-159540/14.

XX N-PSDB: V61529.

XX Chimeric protein of leukotoxin and gonadotropin releasing hormone -

XX useful for, e.g. preparation of vaccines for reduction of incidence

XX of mammary tumours in mammals

XX Disclosure; Figure 1B; 118pp; English.

XX The present sequence represents a recombinantly produced or chemically

XX synthesised gonadotropin releasing hormone-2 (GnRH-2) polypeptide, which

XX contains four copies of the GnRH decapeptide (V61528) and triplet amino

XX acid spacers between each of these sequences. This decapeptide is

XX secreted naturally by the hypothalamus which controls release of both the

XX luteinising hormone (LH) and the follicle stimulating hormone (FSH) in

XX vertebrates. This sequence, as compared to the native peptide, has been

XX found to have an N-terminal Gln rather than a pyroglu residue, and also

XX contains substitutions at amino acid residues 15 and 41, whereby His is

XX replaced by Asp. This produces an alternating multimeric GnRH sequence

XX which is highly immunogenic that can be used in the construction of a

XX chimeric protein that comprises a leukotoxin polypeptide, several

XX multimers, and the GnRH sequence. The chimeric protein can be used as a

XX vaccine to help reduce the incidence of mammary tumours in a mammalian

CC Individual.
 XX
 SQ Sequence 49 AA;

Query Match 59.6%; Score 62; DB 19; Length 49;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HNSYGLRPGS 11
 |||||
 DB 2 hwsyglrpgs 11

RESULT 8
 W61542
 ID W61542 standard; Protein; 49 AA.

AC W61542;

DT 27-OCT-1998 (first entry)

DE Peptide hormone GnRH-2 decapeptide (4 copies) fragment.

XX GnRH; gonadotropin releasing hormone peptide hormone; leukotoxin;
 XX immunisation; endogenous molecule; vaccine; ear; immunogen; carrier;
 KM immune response; hormone receptor; cancerous cell; domestic animal;
 KM porcine; bovine; luteinizing hormone; follicle stimulating hormone;
 KM immunocastrate.

OS Synthetic.

PN WO9834639-A1.

XX 13-AUG-1998.

PD 04-FEB-1998; 98WO-CA00059.

XX 05-FEB-1997; 97US-0036883.

XX (BIOS-) BIOSTAR INC.

PI Acres SD, Harland R, Manns JG;

DR WPI: 1998-446952/38.

DR N-PSDB: V45190.

PT Immunisation against endogenous molecules by administering vaccine
 PT to ear - useful to elicit efficient and uniform immune response
 PT against e.g. gonadotropin releasing hormone to. immunocastrate pigs
 PT and cattle

XX Example 1; Fig 1B; 61pp; English.

CC This represents the amino acid sequence of the gonadotropin releasing
 CC hormone (GnRH-2) decapeptide fragment used in the chimeric leukotoxin-
 CC GnRH polypeptide gene fusions. This is used to exemplify the method of
 CC invention of immunisation against endogenous molecules by administering
 CC a vaccine which comprises an immunogen and a carrier to the ear of the
 CC mammal. The method is useful for eliciting an efficient and uniform
 CC immune response to block or suppress the activity of an endogenous
 CC hormone, hormone receptor, agonist or antagonist in a vaccinated subject,
 CC or to elicit an immune response against a targeted endogenous cell type
 CC (e.g. a cancerous or otherwise diseased cell). It is especially useful
 CC to reduce the levels of GnRH in domestic animals, especially in porcine
 CC or bovine species. The use of GnRH immunogens in the vaccine reduces the
 CC levels of luteinizing hormone and follicle stimulating hormone and helps
 CC in immunocastrating the animal. Administration of vaccine compositions to
 CC the ear instead of intramuscular administration into the neck increases
 CC the efficiency of vaccination of mammals against endogenous immunogens,
 CC and may increase uniformity of vaccine presentation since the ear is
 CC relatively uniform from animal to animal.

SQ Sequence 49 AA;

Query Match 59.6%; Score 62; DB 19; Length 49;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HNSYGLRPGS 11
 |||||
 DB 2 hwsyglrpgs 11

RESULT 9
 Y58363
 ID Y58363 standard; Protein; 49 AA.

AC Y58363;

DT 27-MAR-2000 (first entry)

DE Four-copy gonadotropin-releasing hormone (GnRH) multimer.

XX GnRH multimer; gonadotropin-releasing hormone; immunosterilisation;
 XX immunocastration; vaccine; feline; canine; equine; cervine; ds
 OS Mammalia.
 OS Synthetic.

PN WO962545-A2.

PD 09-DEC-1999.

XX 28-MAY-1999; 99WO-CA00493.

XX 04-JUN-1998; 98US-0088024.

XX 06-MAY-1999; 99US-0306689.

XX (BIOS-) BIOSTAR INC.

PI Robbins SC;

DR WPI: 2000-086857/07.

DR N-PSDB: Z55702.

PT Hormone immunogens, analogues or antibodies used to manufacture
 PT vaccines for suppression of reproductive behavior and fertility in
 PT vertebrates -
 PT
 PS Claim 5; Fig 5B; 88pp; English.

CC This sequence represents a four-copy gonadotropin-releasing
 CC hormone (GnRH) multimer, where the second and fourth GnRH sequence
 CC have a His to Asp substitution at position 2 of the GnRH sequence.
 CC The invention relates to GnRH immunogens, analogues or antibodies
 CC that cross-react with endogenous GnRH of a vertebrate. A specifically
 CC claimed immunogenic fusion protein (Y58361) comprises, in the N to
 CC C-terminal direction, a synthetic peptide sequence (Y58364), an eight
 CC copy GnRH multimer (composed of two copies of the 4xGnRH multimer
 CC sequence of Y58363), the LRT protein (which functions as a carrier
 CC protein), and a second eight copy GnRH multimer. The fusion protein may
 CC be used in a vaccine composition for prepubertal administration to a
 CC vertebrate subject to result in prolonged suppression of reproductive
 CC behaviour and/or fertility. GnRH immunogens, analogues or antibodies are
 CC used to manufacture a composition or vaccine for immunosterilisation or
 CC immunocastration of feline, canine, equine or cervine subjects.
 CC The vaccines are used to suppress reproductive behaviour and/or
 CC fertility for at least 10 months. The prepubertal administration
 CC results in a prolonged, long-term suppression of testicular development
 CC and/or function in males, or a prolonged, long-term suppression of
 CC ovarian development and/or function in females. The methods provide a
 CC viable and desirable alternative to surgical forms of sterilisation that
 CC are currently used.

SQ Sequence 49 AA:

Query Match 59.6%; Score 62; DB 21; Length 49;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 11
|||
Db 2 hwsyglrpgs 11

RESULT 10
Y58135 Y58135 standard; Protein; 49 AA.

XX AC Y58135;

XX DT 07-MAR-2000 (first entry)

XX DE GnRH analogue multimer, containing four copies of the GnRH analogue.

KW Gonadotropin releasing hormone; GnRH; leukotoxin; LKT; fusion protein;
KW antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;
KW non-androgenic; steroid; reduction; weight gain; muscle distribution;
KW fat distribution; male pattern; boar taint; flavour; impairment;
KW reliable; immunocastration; meat production.

XX Synthetic.
OS Mammalia.

XX PN WO9956771-A2.

XX PD 11-NOV-1999.

XX PF 05-MAY-1999; 99WO-CA00360.

XX PR 05-MAY-1998; 98US-0084217.

XX PA (BIOS-) BIOSTAR INC.

XX PI Manns JG, Acres SD, Harland R;

XX DR WPI: 2000-062125/05.

XX DR N-PSDB; 246402.

XX PT Production of uncastrated male food animals using vaccines -

XX PS Example 1; Fig 2B; 87pp; English.

XX This sequence represents four copies of a gonadotropin
CC releasing hormone (GnRH) analogue, DNA encoding which was
CC used in the construction of a chimeric GnRH-leukotoxin (LKT)
CC fusion gene (Z46400). This fusion gene encodes a GnRH-LKT fusion
CC protein which may be used as a vaccine. The LKT portion of the protein
CC acts to enhance the immunogenicity of the GnRH portion. The invention
CC relates to a method of using two GnRH immunogen vaccines to produce
CC uncastrated male animals for meat production, one vaccination prior to
CC or during the fattening period to reduce circulating testosterone levels,
CC and the second vaccination about 2-8 weeks before slaughter to
CC substantially reduce androgenic and/or non-androgenic steroids. The
CC invention is used to produce food animals that exhibit the weight gain
CC and muscle/fat distribution of male animals without the problems
CC associated with male animals. Such problems include "boar taint", a
CC urine-like odour found in cooked meat of uncastrated pigs which is
CC caused by steroids stored in the tissues, and similar flavour
CC impairments in the meat of other intact male animals. The invention is
CC more reliable than prior art immunocastration techniques.

SQ Sequence 49 AA:

Query Match 59.6%; Score 62; DB 21; Length 49;

Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 11
|||
Db 2 hwsyglrpgs 11

RESULT 11
R11178 R11178 standard; Protein; 257 AA.

XX AC R11178;

XX DT 22-MAY-1991 (first entry)

XX DE Plasmid pBTAT735-encoded TrATp-LHRH fusion protein.

KW TrAT protein; Leutinizing hormone releasing hormone; fusion protein;
KW immunological castration.

XX FH Key location/Qualifiers

XX FT Peptide 1..20 /label= TrATp signal

XX FT Peptide 103..112 /label= LHRH analogue

XX FT Protein 21..257 /label= TrATp-LHRH fusion

XX FN WO9102799-A.

XX PD 07-MAR-1991.

XX PF 24-AUG-1990; 90WO-AU00373.

XX PR 25-AUG-1989; 89AU-0005979.

XX PA (BIOT-) BIOTECHN AUST PTY L.

XX PI Russell-Jones GJ, Stewart AG, Tsonis CG;

XX DR WPI: 1991-087282/12.

XX DR N-PSDB; Q10999.

XX PT Fusion proteins comprising LHRH analogue and TrATp (analogue) -
XX PT useful in vaccine for inhibition or control of reproduction in
XX PT vertebrates, esp. domestic animals

XX PS Example 1; Fig 2 and 3; 53pp; English.

XX Plasmid pBTAT735 is a TrATp-LHRH analogue fusion in which the LHRH
CC analogue has been inserted between amino acids 101 and 102 of TrATp
CC (Ogata R.T. et al., (1982) J. Bacteriol. 151:819-827). The synthetic
CC LHRH analogue was inserted via a linker which provides a unique new
CC SmaI site located between codons such that the LHRH is inserted in
CC frame. The fusion protein can be used in vaccines for the
CC inhibition or control of reproduction in vertebrates, where the
CC TrATp acts as an adjuvant.
CC See also Q10995, Q10997-8, Q11000, Q11014-Q11021.

SQ Sequence 257 AA:

Query Match 59.6%; Score 62; DB 12; Length 257;
Best Local Similarity 90.9%; Pred. No. 0.96;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSS 12
|||
Db 104 hwsyglrpgss 114

RESULT 12

R1183
ID R1183 standard; Protein: 257 AA.
XX
AC R1183;
XX
DT 22-MAY-1991 (first entry)
XX
DE Plasmid pBTA740-encoded Tratp-LHRH fusion protein.
XX
KW Trat protein; leutinizing hormone releasing hormone; fusion protein;
KW Immunological castration.
XX
FH Key Location/Qualifiers
FT Peptide 1..20
FT /label= Tratp signal
FT Peptide 238..247
FT /label= LHRH analogue
FT Protein 21..257
FT /label= Tratp-LHRH fusion
XX
PN W09102799-A.
XX
PD 07-MAR-1991.
XX
PF 24-AUG-1990; 90MO-AU00373.
XX
PR 25-AUG-1989; 89AU-0005979.
XX
PA (BIOT-) BIOTECHN AUST PTY L.
XX
PI Russell-Jones GJ, Stewart AG, Tsouis CG;
XX
DR WPI; 1991-087282/12.
DR N-PSDB; Q11017.
XX
XX Fusion proteins comprising LHRH analogue and Tratp (analogue) -
PT useful in vaccine for inhibition or control of reproduction in
PT vertebrates, esp. domestic animals
XX
PS Example 1; Fig 2 and 3; 53pp; English.
XX
CC Plasmid pBTA740 is a Tratp-LHRH analogue fusion in which the LHRH
CC analogue has been inserted between amino acids 236 and 237 of Tratp
CC (Ogata R.T. et al., (1982) J Bacteriol. 151:819-827). The synthetic
CC LHRH analogue was inserted via a linker which provides a unique new
CC SmaI site located between codons such that the LHRH is inserted in
CC frame. The fusion protein can be used in vaccines
CC for the inhibition or control of reproduction in vertebrates, where
CC the Tratp acts as an adjuvant.
CC See also Q10995, Q10997-Q11000, Q11014-6, Q11018-Q11021.
XX
SQ Sequence 257 AA;

Query Match 59.6%; Score 62; DB 12; Length 257;
Best Local Similarity 90.9%; Pred. No. 0.96;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSS 12
DB 239 hwsyglrpgss 249
|||||||
|

RESULT 13
ID W03943
W03943 standard; Protein: 544 AA.
XX
AC W03943;
XX
DT 20-NOV-1996 (first entry)
XX
DE LKT-GnRH protein fusion from PCB111.
XX

KW Leukotoxin; LKT; gonadotropin-releasing hormone; GnRH;
KW fusion protein; immunogen; vaccine; fertility control;
KW contraceptive; sterilisation.
XX
XX OS Chimeric Pasteurella haemolytica A1 strain B122;
XX Chimeric synthetic.
XX
FH Key Location/Qualifiers
FT Domain 1..493
FT /label= LKT
FT Domain 494..544
FT /label= GnRH_repeat_domain
XX
PN W09624675-A1.
XX
PD 15-AUG-1996.
XX
PF 24-JAN-1996; 96MO-CA00049.
XX
PR 10-FEB-1995; 95US-0387156.
XX
PA (UYSA-) UNIV SASKATCHEWAN.
XX
PI Manns JG, Potter AA;
XX
DR WPI; 1996-384447/38.
DR N-PSDB; T37177.
XX
XX Gonadotropin-releasing hormone multimer fusion proteins - with
PT leukotoxin polypeptide for increased immunogenicity, useful in
PT antifertility vaccine prodn.
XX
PS Claim 8; Fig 7A-7E; 87pp; English.
XX
XX A chimeric protein (W03943) is composed of a fusion between
CC a truncated leukotoxin (LKT-111) from Pasteurella haemolytica
CC and a 4-copy gonadoliberin-releasing hormone (GnRH) repeat
CC sequence (see also W03944). It is the product of a chimeric
CC gene (T37177) produced by deleting an approx. 1300 bp sequence
CC from PCB113 (see also T37176) coding for amino acids 352-784
CC of LKT-352. Recombinant plasmid PCB111 (LKT 111:4 copy GnRH,
CC ATCC 69748) was obtd. Escherichia coli transformants produced
CC the chimeric protein, which is useful as a vaccine for fertility
CC control, esp. immunological sterilisation of domestic or farm
CC animals.
XX
SQ Sequence 544 AA;

Query Match 59.6%; Score 62; DB 17; Length 544;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGS 11
DB 495 hwsyglrpgs 504
|||||||
|

RESULT 14
ID W79570
W79570 standard; Protein: 544 AA.
XX
AC W79570;
XX
DT 24-DEC-1998 (first entry)
XX
DE LKT-GnRH chimeric protein.
XX
XX Chimera; PCB111; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;
KW cytotoxic activity; antigen presentation; immune response; vaccine;
KW tumour.
XX
OS Synthetic.

XX MO9806848-A1.
 PN 19-FEB-1998.
 PD 08-AUG-1997; 97WO-CA00559.
 PE 09-AUG-1996; 96US-0694865.
 PR (UYSA-) UNIV SASKATCHEWAN.
 PA Manns JG, Potter AA;
 PI WPI; 1998-159540/14.
 DR N-PSDB; V61532.
 XX Chimeric protein of leukotoxin and gonadotropin releasing hormone -
 PT useful for, e.g. preparation of vaccines for reduction of incidence
 of mammary tumours in mammals
 PS Disclosure; Figure 7.1-5; 118pp; English.
 XX The present sequence represents the LKT-GnRH chimeric protein from
 CC PCB111. This plasmid contains the LKT 111 polypeptide fused to
 CC four copies of the GnRH peptide. This chimera lacks cytotoxic activity
 CC which enables there to be an increase in antigen presentation and thus an
 CC optimal immune response. The removal of this region also enables the
 CC truncated LKT to be expressed at much higher levels and allows the amount
 CC of antigen administered to be reduced. This chimeric protein comprises a
 CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The
 CC chimeric protein can be used as a vaccine to help reduce the incidence of
 CC mammary tumours in a mammalian individual.
 SQ Sequence 544 AA;

Query Match 59.6%; Score 62; DB 19; Length 544;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 11
 |||||
 DB 495 hwsyglrpgs 504

RESULT 15
 W79573
 ID W79573 standard; Protein: 695 AA.
 XX W79573;
 AC 24-DEC-1998 (first entry)
 DT LKT-GnRH chimeric protein.
 XX
 DE Chimera; PCB122; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;
 KW cytotoxic activity; antigen presentation; immune response; vaccine;
 tumour.
 XX Synthetic.
 OS
 XX MO9806848-A1.
 PN 19-FEB-1998.
 PD 08-AUG-1997; 97WO-CA00559.
 PE 09-AUG-1996; 96US-0694865.
 PR (UYSA-) UNIV SASKATCHEWAN.
 PA Manns JG, Potter AA;
 PI
 XX

DR WPI; 1998-159540/14.
 DR N-PSDB; V61535
 XX Chimeric protein of leukotoxin and gonadotropin releasing hormone -
 PT useful for, e.g. preparation of vaccines for reduction of incidence
 of mammary tumours in mammals
 PS Claim 9; Figure 9.1-6; 118pp; English.
 XX The present sequence represents the LKT-GnRH chimeric protein from
 CC PCB122. This plasmid contains the LKT 111 polypeptide fused to sixteen
 CC copies of the GnRH peptide. In the pattern of: 8 copies of GnRH-LKT 111-8
 CC copies of GnRH. This chimera lacks cytotoxic activity which enables
 CC there to be an increase in antigen presentation and thus an optimal
 CC immune response. The removal of this region also enables the truncated
 CC LKT to be expressed at much higher levels and allows the amount of
 CC antigen administered to be reduced. This chimeric protein comprises a
 CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The
 CC chimeric protein can be used as a vaccine to help reduce the incidence of
 CC mammary tumours in a mammalian individual.
 SQ Sequence 695 AA;

Query Match 59.6%; Score 62; DB 19; Length 695;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 11
 |||||
 DB 10 hwsyglrpgs 19

Search completed: March 2, 2001, 10:53:46
 Job time: 911 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:59:03 ; Search time 201.99 Seconds
(without alignments)
12.920 Million cell updates/sec

Title: US-09-306-689-4

Perfect score: 104

Sequence: 1 XHWSYGLRPGSSPPPC 17

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 957798 segs, 153513742 residues

Total number of hits satisfying chosen parameters: 957798

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA:*
1: /cgn2_6/ptodata/2/paa/PCTUS.COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06.COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07.COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081.COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082.COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083.COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084.COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085.COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086.COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087.COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088.COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089.COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090.COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091.COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092.COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093.COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094.COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095.COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096.COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097.COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098.COMB.pep.*
23: /cgn2_6/ptodata/2/paa/PCT_NEW.COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US06_NEW.COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US07_NEW.COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US08_NEW.COMB.pep.*
27: /cgn2_6/ptodata/2/paa/US09_NEW.COMB.pep.*
28: /cgn2_6/ptodata/2/paa/US06_NEW.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	103	99.0	17	1 PCT-US95-01225-6	Sequence 6, Appl
2	103	99.0	17	8 US-08-478-546B-6	Sequence 6, Appl
3	103	99.0	17	17 US-09-305-924-4	Sequence 4, Appl
4	103	99.0	17	17 US-09-306-689-4	Sequence 4, Appl
5	83.5	80.3	16	1 PCT-US95-01225-2	Sequence 2, Appl

6	83.5	80.3	16	8 US-08-478-546B-2	Sequence 2, Appl
7	83.5	80.3	16	17 US-09-305-924-5	Sequence 5, Appl
8	83.5	80.3	16	17 US-09-306-689-5	Sequence 5, Appl
9	65.5	63.0	411	19 US-09-506-078-23	Sequence 23, Appl
10	65.5	63.0	411	22 US-60-120-454-23	Sequence 23, Appl
11	65.5	63.0	442	19 US-09-506-078-27	Sequence 27, Appl
12	65.5	63.0	442	19 US-09-506-078-27	Sequence 27, Appl
13	65.5	63.0	442	22 US-60-120-454-27	Sequence 27, Appl
14	62	59.6	17	5 US-08-160-882-18	Sequence 18, Appl
15	62	59.6	18	5 US-08-160-882-28	Sequence 28, Appl
16	62	59.6	49	14 US-09-019-010-4	Sequence 4, Appl
17	62	59.6	49	16 US-09-249-447A-7	Sequence 7, Appl
18	62	59.6	49	17 US-09-305-924-11	Sequence 11, Appl
19	62	59.6	49	17 US-09-306-689-11	Sequence 11, Appl
20	62	59.6	49	17 US-09-383-912-4	Sequence 4, Appl
21	62	59.6	544	17 US-09-383-912-10	Sequence 10, Appl
22	62	59.6	635	17 US-09-305-924-13	Sequence 13, Appl
23	62	59.6	635	17 US-09-306-689-13	Sequence 13, Appl
24	62	59.6	699	17 US-09-383-912-16	Sequence 16, Appl
25	60	57.7	977	17 US-09-383-912-8	Sequence 8, Appl
26	59	56.7	372	16 US-09-252-991A-20108	Sequence 20108, A
27	59	56.7	14	5 US-08-160-882-22	Sequence 22, Appl
28	59	56.7	14	5 US-08-160-882-26	Sequence 26, Appl
29	59	56.7	14	5 US-08-160-882-26	Sequence 26, Appl
30	59	56.7	18	1 PCT-US99-11219-1148	Sequence 1148, Ap
31	59	56.7	18	14 US-09-082-279-1148	Sequence 1148, Ap
32	59	56.7	18	14 US-09-082-279-1148	Sequence 1148, Ap
33	59	56.7	18	14 US-09-082-279-1148	Sequence 1148, Ap
34	59	56.7	18	17 US-09-315-304-1148	Sequence 1148, Ap
35	59	56.7	18	17 US-09-350-641-1148	Sequence 1148, Ap
36	59	56.7	18	23 PCT-US01-03968-1148	Sequence 1148, Ap
37	59	56.7	24	5 US-08-160-882-43	Sequence 43, Appl
38	59	56.7	44	5 US-08-160-882-45	Sequence 45, Appl
39	59	56.7	84	5 US-08-160-882-47	Sequence 47, Appl
40	58	55.8	9	20 US-09-657-276-153	Sequence 153, App
41	58	55.8	10	1 PCT-US00-22121-28	Sequence 28, Appl
42	58	55.8	10	1 PCT-US00-22121-32	Sequence 32, Appl
43	58	55.8	10	1 PCT-US94-0483A-1	Sequence 1, Appl
44	58	55.8	10	1 PCT-US94-12765-1	Sequence 4, Appl
45	58	55.8	10	27 US-09-412-558-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
PCT-US95-01225-6
; Sequence 6, Application PC/US9501225
; GENERAL INFORMATION:
; APPLICANT: Scimens, Stephen
; TITLE OF INVENTION: Immunogens Against Gonadotropin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Dimitrios T. Drivas, Esq.
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01225
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Drivas Esq., Dimitrios T.
; REGISTRATION NUMBER: 32,218

REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8286
TELEFAX: 212-354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label=pglu
OTHER INFORMATION: /note="pyroglutamic acid (5-oxoproline)"
FEATURE:
NAME/KEY: Region
LOCATION: 1..10
OTHER INFORMATION: /note="immunomimic"
FEATURE:
NAME/KEY: Region
LOCATION: 11..17
OTHER INFORMATION: /note="spacer"
PCT-US95-01225-6

Query Match 99.0%; Score 103; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSSPPPC 17
Db 2 HWSYGLRPGSSPPPC 17

RESULT 2
US-08-478-546B-6
Sequence 6, Application US/08478546B
GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
APPLICANT: Scibbenski, Robert
TITLE OF INVENTION: Methods for the Treatment of Hormone-Dependent
TITLE OF INVENTION: Tumors with Immunogens against Gonadotropin Releasing Hormone
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dimitrios T. Drivas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,546B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/188,223
FILING DATE: 27-JAN-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218
REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8286
TELEFAX: 212-354-8113
INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified site
LOCATION: 1
OTHER INFORMATION: /note= Xaa
OTHER INFORMATION: /note="pyroglutamic acid"
FEATURE:
NAME/KEY: Region
LOCATION: 1..10
OTHER INFORMATION: /note="immunomimic"
FEATURE:
NAME/KEY: Region
LOCATION: 11..17
OTHER INFORMATION: /note="spacer"
US-08-478-546B-6

Query Match 99.0%; Score 103; DB 8; Length 17;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSSPPPC 17
Db 2 HWSYGLRPGSSPPPC 17

RESULT 3
US-09-305-924-4
Sequence 4, Application US/09305924A
GENERAL INFORMATION:
APPLICANT: Jack G. Manns
APPLICANT: Stephen D. Acres
APPLICANT: Richard Harland
TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
FILE REFERENCE: 9001-0048
CURRENT APPLICATION NUMBER: US/09/305,924A
EARLIER APPLICATION NUMBER: US 60/084,217
EARLIER FILING DATE: 1998-05-05
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 4
LENGTH: 17
TYPE: PRT
ORGANISM: GnRH
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (1)
OTHER INFORMATION: Xaa is pyroglutamic acid
US-09-305-924-4

Query Match 99.0%; Score 103; DB 17; Length 17;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HWSYGLRPGSSPPPC 17
Db 2 HWSYGLRPGSSPPPC 17

RESULT 4
US-09-306-689-4
Sequence 4, Application US/09306689B
GENERAL INFORMATION:
APPLICANT: Robbins, Sarah C.
TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
TITLE OF INVENTION: ANIMALS
FILE REFERENCE: 9001-0047

;; CURRENT APPLICATION NUMBER: US/09/306,689B
;; CURRENT FILING DATE: 1999-05-06
;; EARLIER APPLICATION NUMBER: US 60/088,024
;; EARLIER FILING DATE: 1998-06-04
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 4
;; LENGTH: 17
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; NAME/KEY: MOD_RES
;; LOCATION: (1)
;; OTHER INFORMATION: Xaa is pyroglutamic acid
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Sequence
;; OTHER INFORMATION: defines GNRH analogue
;; US-09-306-689-4

Query Match 99.0%; Score 103; DB 17; Length 17;
Best Local Similarity 100.0%; Pred. No. 1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSSPPPC 17
Db 2 HWSYGLRPGSSPPPC 17

RESULT 5
PCT-US95-01225-2
; Sequence 2, Application pc/TUS9501225
; GENERAL INFORMATION:
; APPLICANT: Grimes, Stephen
; APPLICANT: Scibienski, Robert
; TITLE OF INVENTION: Immunogens Against Gonadotropin
; TITLE OF INVENTION: Releasing Hormone
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Dimitrios T. Drivas, Esq.
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01225
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Drivas Esq., Dimitrios T.
; REGISTRATION NUMBER: 32,218
; REFERENCE/DOCKET NUMBER: 1102865-300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-819-8286
; TELEFAX: 212-354-8113
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..10
; OTHER INFORMATION: /note= "immunomimic"
; FEATURE:
; NAME/KEY: Region

;; NAME/KEY: Region
;; LOCATION: 11..16
;; OTHER INFORMATION: /note= "spacer"
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 1
;; OTHER INFORMATION: /label= pglu
;; OTHER INFORMATION: /note= "pyroglutamic acid (5-oxoproline)"
;; PCT-US95-01225-2

Query Match 80.3%; Score 83.5; DB 1; Length 16;
Best Local Similarity 87.5%; Pred. No. 0.0019;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 2 HWSYGLRPGSSPPPC 17
Db 2 HWSYGLRPG-RPPPC 16

RESULT 6
US-08-478-546B-2
; Sequence 2, Application US/08478546B
; GENERAL INFORMATION:
; APPLICANT: Grimes, Stephen
; APPLICANT: Scibienski, Robert
; TITLE OF INVENTION: Methods for the Treatment of Hormone-Dependent
; TITLE OF INVENTION: Tumors with Immunogens against Gonadotropin Releasing Hormo
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Dimitrios T. Drivas, Esq.
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,546B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/188,223
; FILING DATE: 27-JAN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Drivas Esq., Dimitrios T.
; REGISTRATION NUMBER: 32,218
; REFERENCE/DOCKET NUMBER: 1102865-300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-819-8286
; TELEFAX: 212-354-8113
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified site
; LOCATION: 1
; OTHER INFORMATION: /note= Xaa
; OTHER INFORMATION: /note= "pyroglutamic acid"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..10
; OTHER INFORMATION: /note= "immunomimic"
; FEATURE:
; NAME/KEY: Region

LOCATION: 11..16
OTHER INFORMATION: /note= 'spacer'
US-08-478-546B-2

Query Match 80.3%; Score 83.5; DB 8; Length 16;
Best Local Similarity 87.5%; Pred. No. 0.0019;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
OY 2 HWSYGLRPGSSPPPC 17
Db 2 HWSYGLRPG-RPPPC 16

RESULT 7
US-09-305-924-5
Sequence 5, Application US/09305524A
GENERAL INFORMATION:
APPLICANT: Jack G. Manns
APPLICANT: Stephen D. Acres
APPLICANT: Richard Harland
TITLE OF INVENTION: METHODS OF FAISING ANIMALS FOR MEAT PRODUCTION
FILE REFERENCE: 9001-0048
CURRENT APPLICATION NUMBER: US/09/305,924A
EARLIER FILING DATE: 1999-05-05
EARLIER APPLICATION NUMBER: US 60/084,217
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 16
TYPE: PRT
ORGANISM: GnRH
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (1)
OTHER INFORMATION: Xaa is pyroglutamic acid
US-09-305-924-5

Query Match 80.3%; Score 83.5; DB 17; Length 16;
Best Local Similarity 87.5%; Pred. No. 0.0019;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
OY 2 HWSYGLRPGSSPPPC 17
Db 2 HWSYGLRPG-RPPPC 16

RESULT 8
US-09-306-689-5
Sequence 5, Application US/09306689B
GENERAL INFORMATION:
APPLICANT: Robbins, Sarah C.
TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
FILE REFERENCE: 9001-0047
CURRENT APPLICATION NUMBER: US/09/306,689B
EARLIER FILING DATE: 1999-05-06
EARLIER APPLICATION NUMBER: US 60/088,024
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 16
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (1)
OTHER INFORMATION: Xaa is pyroglutamic acid
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Sequence

OTHER INFORMATION: defines GnRH analogue
US-09-306-689-5

Query Match 80.3%; Score 83.5; DB 17; Length 16;
Best Local Similarity 87.5%; Pred. No. 0.0019;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
OY 2 HWSYGLRPGSSPPPC 17
Db 2 HWSYGLRPG-RPPPC 16

RESULT 9
US-09-506-078-23
Sequence 23, Application US/09506078
GENERAL INFORMATION:
APPLICANT: Pfizer Products Inc
TITLE OF INVENTION: FUSION PROTEINS COMPRISING CARRIERS THAT CAN INDUCE A
FILE REFERENCE: PCI0202A
CURRENT APPLICATION NUMBER: US/09/506,078
EARLIER FILING DATE: 2000-02-16
EARLIER APPLICATION NUMBER: N/A
EARLIER FILING DATE: 1999-02-17
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 23
LENGTH: 411
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: 4GnRH-tm9D
US-09-506-078-23

Query Match 63.0%; Score 65.5; DB 19; Length 411;
Best Local Similarity 81.2%; Pred. No. 4.5;
Matches 13; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
OY 2 HWSYGLRPGS-SPPP 16
Db 50 HWSYGLRPGSMSPPT 65

RESULT 10
US-60-120-454-23
Sequence 23, Application US/60120454
GENERAL INFORMATION:
APPLICANT: Campos, Manuel
APPLICANT: Yule, Terecila D
APPLICANT: Martindog, Serge
APPLICANT: Ditschil, Becky A
TITLE OF INVENTION: FUSION PROTEINS COMPRISING CARRIERS THAT CAN INDUCE A
FILE REFERENCE: PCI0202
CURRENT APPLICATION NUMBER: US/60/120,454
EARLIER FILING DATE: 1999-02-17
EARLIER APPLICATION NUMBER: N/A
NUMBER OF SEQ ID NOS: 46
SOFTWARE: PatentIn Ver. 2.0 - beta
SEQ ID NO 23
LENGTH: 411
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: 4GnRH-tm9D
US-60-120-454-23

Query Match 63.0%; Score 65.5; DB 22; Length 411;
Best Local Similarity 81.2%; Pred. No. 4.5;
Matches 13; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2 HWSYGLRPGS-SPPPP 16
DB 50 HWSYGLRPGSMSLPTP 65

RESULT 11

US-09-506-078-27
; Sequence 27, Application US/09506078
; GENERAL INFORMATION:
; APPLICANT: Pfizer Products Inc
; TITLE OF INVENTION: FUSION PROTEINS COMPRISING CARRIERS THAT CAN INDUCE A
; FILE REFERENCE: PC10202A
; CURRENT APPLICATION NUMBER: US/09/506,078
; CURRENT FILING DATE: 2000-02-16
; EARLIER APPLICATION NUMBER: N/A
; EARLIER FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
US-09-506-078-27

Query Match 63.0%; Score 65.5; DB 19; Length 442;
Best Local Similarity 81.2%; Pred. No. 4.8;
Matches 13; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2 HWSYGLRPGS-SPPPP 16
DB 50 HWSYGLRPGSMSLPTP 65

RESULT 12

US-60-120-454-27
; Sequence 27, Application US/60120454
; GENERAL INFORMATION:
; APPLICANT: Campos, Manuel
; APPLICANT: Yule, Terecila D
; APPLICANT: Martind, Serge
; APPLICANT: Dirtschl, Becky A
; TITLE OF INVENTION: FUSION PROTEINS COMPRISING CARRIERS THAT CAN INDUCE A
; FILE REFERENCE: PC10202
; CURRENT APPLICATION NUMBER: US/60/120,454
; CURRENT FILING DATE: 1999-02-17
; EARLIER APPLICATION NUMBER: N/A
; EARLIER FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 27
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: 4GnRH-tmpd-4GnRH encoded by pOE-GnRH:gd:GnRH
US-60-120-454-27

Query Match 63.0%; Score 65.5; DB 22; Length 442;
Best Local Similarity 81.2%; Pred. No. 4.8;
Matches 13; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2 HWSYGLRPGS-SPPPP 16
DB 50 HWSYGLRPGSMSLPTP 65

RESULT 13

US-08-160-882-18
; Sequence 18, Application US/08160882
; GENERAL INFORMATION:
; APPLICANT: Russel-Jones, Gregory J.
; APPLICANT: Stewart, Andrew G.
; APPLICANT: Tsouis, Con G.
; TITLE OF INVENTION: Triat/Neuropeptide-Y Fusion Proteins
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/160,882
; FILING DATE: 03-DEC-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/690,983
; FILING DATE: 25-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU90/00373
; FILING DATE: 24-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU P3 5979
; FILING DATE: 25-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 60042/119/BINU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-160-882-18

Query Match 59.6%; Score 62; DB 5; Length 17;
Best Local Similarity 90.9%; Pred. No. 0.7;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPGSS 12
DB 5 HWSYGLRPGS 15

RESULT 14

US-08-160-882-28
; Sequence 28, Application US/08160882
; GENERAL INFORMATION:
; APPLICANT: Russel-Jones, Gregory J.
; APPLICANT: Stewart, Andrew G.
; APPLICANT: Tsouis, Con G.
; TITLE OF INVENTION: Triat/Neuropeptide-Y Fusion Proteins
; NUMBER OF SEQUENCES: 55

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Foley & Lardner
;; STREET: 3000 K Street, N.W., Suite 500
;; CITY: Washington, D.C.
;; COUNTRY: USA
;; ZIP: 20007-5109
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/160,882
;; FILING DATE: 03-DEC-1993
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/690,983
;; FILING DATE: 25-JUN-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/AU90/00373
;; FILING DATE: 24-AUG-1990
;; PRIOR APPLICATION DATA:
;; FILING DATE: 25-AUG-1989
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BENT, Stephen A.
;; REGISTRATION NUMBER: 29,768
;; REFERENCE/DOCKET NUMBER: 60042/119/BIU
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)672-5300
;; TELEFAX: (202)672-5399
;; TELEX: 904136
;;
;; INFORMATION FOR SEQ ID NO: 28:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 18 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
US-08-160-882-28

Query Match 59.6%; Score 62; DB 5; Length 18;
Best Local Similarity 90.9%; Pred. NO. 0.73;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 HWSYGLRPGSS 12
| | | | | | | | | |
Db 6 HWSYGLRPGGS 16

RESULT 15
US-09-019-010-4
;; Sequence 4, Application US/09019010
;; GENERAL INFORMATION:
;; APPLICANT: HARLAND, RICHARD
;; APPLICANT: MANN, JOHN G.
;; APPLICANT: ACRES, STEPHEN D.
;; TITLE OF INVENTION: IMMUNIZATION AGAINST ENDOGENOUS
;; TITLE OF INVENTION: MOLECULES
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ROHINS & ASSOCIATES
;; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
;; CITY: MENLO PARK
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94025
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/09/019,010
;; FILING DATE: 05-FEB-1998
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 60/036,883
;; FILING DATE: 05-FEB-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: MCCracken, THOMAS P.
;; REGISTRATION NUMBER: 38,548
;; REFERENCE/DOCKET NUMBER: 9001-0035
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (650) 325-7812
;; TELEFAX: (650) 325-7823
;;
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 49 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
US-09-019-010-4

Query Match 59.6%; Score 62; DB 14; Length 49;
Best Local Similarity 100.0%; Pred. NO. 1.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGS 11
| | | | | | | | | |
Db 2 HWSYGLRPGS 11

Search completed: March 2, 2001, 10:59:03
Job time: 373 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:59:03 ; Search time 201.99 Seconds
(without alignments)
12.160 Million cell updates/sec

Title: US-09-306-689-5

Perfect score: 101

Sequence: 1 XHMSYGLRPGRRPPPC 16

Scoring table:

BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 957798 seqs, 153513742 residues

Total number of hits satisfying chosen parameters: 957798

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_AA:*

1:	/cgn2_6/ptodata/2/paa/PCRUS.COMB.pep.*
2:	/cgn2_6/ptodata/2/paa/US06.COMB.pep.*
3:	/cgn2_6/ptodata/2/paa/US07.COMB.pep.*
4:	/cgn2_6/ptodata/2/paa/US08.COMB.pep.*
5:	/cgn2_6/ptodata/2/paa/US081.COMB.pep.*
6:	/cgn2_6/ptodata/2/paa/US082.COMB.pep.*
7:	/cgn2_6/ptodata/2/paa/US083.COMB.pep.*
8:	/cgn2_6/ptodata/2/paa/US084.COMB.pep.*
9:	/cgn2_6/ptodata/2/paa/US085.COMB.pep.*
10:	/cgn2_6/ptodata/2/paa/US086.COMB.pep.*
11:	/cgn2_6/ptodata/2/paa/US087.COMB.pep.*
12:	/cgn2_6/ptodata/2/paa/US088.COMB.pep.*
13:	/cgn2_6/ptodata/2/paa/US089.COMB.pep.*
14:	/cgn2_6/ptodata/2/paa/US090.COMB.pep.*
15:	/cgn2_6/ptodata/2/paa/US091.COMB.pep.*
16:	/cgn2_6/ptodata/2/paa/US092.COMB.pep.*
17:	/cgn2_6/ptodata/2/paa/US093.COMB.pep.*
18:	/cgn2_6/ptodata/2/paa/US094.COMB.pep.*
19:	/cgn2_6/ptodata/2/paa/US095.COMB.pep.*
20:	/cgn2_6/ptodata/2/paa/US096.COMB.pep.*
21:	/cgn2_6/ptodata/2/paa/US097.COMB.pep.*
22:	/cgn2_6/ptodata/2/paa/US098.COMB.pep.*
23:	/cgn2_6/ptodata/2/paa/US099.COMB.pep.*
24:	/cgn2_6/ptodata/2/paa/US100.COMB.pep.*
25:	/cgn2_6/ptodata/2/paa/US101.COMB.pep.*
26:	/cgn2_6/ptodata/2/paa/US102.COMB.pep.*
27:	/cgn2_6/ptodata/2/paa/US103.COMB.pep.*
28:	/cgn2_6/ptodata/2/paa/US104.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	99.0	16	1	PCT-US95-01225-2
2	100	99.0	16	8	US-08-478-546B-2
3	100	99.0	16	17	US-09-305-924-5
4	100	99.0	16	17	US-09-306-689-5
5	83.5	82.7	17	1	PCT-US95-01225-6

6	83.5	82.7	17	8	US-08-478-546B-6	Sequence 6, Appli
7	83.5	82.7	17	17	US-09-305-924-4	Sequence 4, Appli
8	83.5	82.7	17	17	US-09-306-689-4	Sequence 4, Appli
9	60	59.4	16	17	US-09-341-590-103	Sequence 103, App
10	59	58.4	16	17	US-09-341-590-104	Sequence 104, App
11	59	58.4	20	1	PCT-US00-22121-26	Sequence 26, Appl
12	59	58.4	20	3	PCT-US00-22121-30	Sequence 30, Appl
13	59	58.4	20	3	US-07-761-849-10	Sequence 10, Appl
14	59	58.4	20	8	US-08-476-013-10	Sequence 10, Appl
15	59	58.4	20	8	US-08-476-013-12	Sequence 12, Appl
16	59	58.4	20	8	US-08-477-298-10	Sequence 10, Appl
17	59	58.4	20	14	US-09-026-276-26	Sequence 26, Appl
18	59	58.4	20	14	US-09-026-276-30	Sequence 30, Appl
19	59	58.4	20	16	US-09-214-009-1	Sequence 1, Appli
20	59	58.4	20	17	US-09-374-721A-26	Sequence 26, Appl
21	59	58.4	20	17	US-09-374-721A-30	Sequence 30, Appl
22	59	58.4	21	3	US-07-761-849-11	Sequence 11, Appl
23	59	58.4	21	3	US-07-761-849-12	Sequence 12, Appl
24	59	58.4	21	8	US-08-476-013-11	Sequence 11, Appl
25	59	58.4	21	8	US-08-476-013-12	Sequence 12, Appl
26	59	58.4	21	8	US-08-476-013-13	Sequence 13, Appl
27	59	58.4	21	8	US-08-476-013-14	Sequence 14, Appl
28	59	58.4	21	8	US-08-477-298-11	Sequence 11, Appl
29	59	58.4	21	8	US-08-477-298-14	Sequence 14, Appl
30	59	58.4	40	1	PCT-US00-22121-35	Sequence 35, Appl
31	59	58.4	40	14	US-09-026-276-35	Sequence 35, Appl
32	59	58.4	40	17	US-09-374-721A-35	Sequence 35, Appl
33	59	58.4	41	1	PCT-US00-22121-34	Sequence 34, Appl
34	59	58.4	41	14	US-09-026-276-34	Sequence 34, Appl
35	59	58.4	41	19	US-09-374-721A-34	Sequence 23, Appl
36	58.5	57.9	411	19	US-09-506-078-23	Sequence 23, Appl
37	58.5	57.9	411	22	US-60-120-454-23	Sequence 27, Appl
38	58.5	57.9	442	19	US-09-506-078-27	Sequence 27, Appl
39	58.5	57.9	442	20	US-60-120-454-27	Sequence 27, Appl
40	58	57.4	9	22	US-09-657-276-153	Sequence 153, App
41	58	57.4	10	1	PCT-US00-22121-28	Sequence 28, Appl
42	58	57.4	10	1	PCT-US00-22121-32	Sequence 32, Appl
43	58	57.4	10	1	PCT-US00-04832A-1	Sequence 1, Appli
44	58	57.4	10	1	PCT-US94-12763-4	Sequence 4, Appli
45	58	57.4	10	27	US-09-412-558-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
PCT-US95-01225-2
Sequence 2, Application PC/FUS9501225
GENERAL INFORMATION:
APPLICANT: Grimes, Stephen
APPLICANT: Sciablenki, Robert
TITLE OF INVENTION: Immunogens Against Gonadotropin
TITLE OF INVENTION: Releasing Hormone
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dimitrios T. Drivas, Esq.
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/01225
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Drivas Esq., Dimitrios T.
REGISTRATION NUMBER: 32,218

REFERENCE/DOCKET NUMBER: 1102865-300
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-819-8286
TELEFAX: 212-354-8113
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
FEATURE:
NAME/KEY: Region
LOCATION: 1..10
OTHER INFORMATION: /note="immunomimic"
FEATURE:
NAME/KEY: Region
LOCATION: 11..16
OTHER INFORMATION: /note="spacer"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /label="pGlu"
OTHER INFORMATION: /note="pyroglutamic acid (5-oxoproline)"
PCT-US95-01225-2

Query Match 99.0%; Score 100; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPGRRPPPC 16
Db 2 HWSYGLRPGRRPPPC 16

RESULT 2
US-08-478-546B-2
; Sequence 2, Application US/08478546B
; GENERAL INFORMATION:
; APPLICANT: Grimes, Stephen
; APPLICANT: Scibienski, Robert
; TITLE OF INVENTION: Methods for the Treatment of Hormone-Dependent
; TITLES OF INVENTION: Tumors with Immunogens against Gonadotropin Releasing Hormone
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dimitrios T. Drivas, Esq.
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,546B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/188,223
; FILING DATE: 27-JAN-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Drivas Esq., Dimitrios T.
; REGISTRATION NUMBER: 32,218
; REFERENCE/DOCKET NUMBER: 1102865-300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-819-8286
; TELEFAX: 212-354-8113
; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified site
LOCATION: 1
OTHER INFORMATION: /note="Xaa"
OTHER INFORMATION: /note="pyroglutamic acid"
FEATURE:
NAME/KEY: Region
LOCATION: 1..10
OTHER INFORMATION: /note="immunomimic"
FEATURE:
NAME/KEY: Region
LOCATION: 11..16
OTHER INFORMATION: /note="spacer"
US-08-478-546B-2

Query Match 99.0%; Score 100; DB 8; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPGRRPPPC 16
Db 2 HWSYGLRPGRRPPPC 16

RESULT 3
US-09-305-924-5
; Sequence 5, Application US/09305924A
; GENERAL INFORMATION:
; APPLICANT: Stephen D. Manns
; APPLICANT: Richard Harland
; TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
; FILE REFERENCE: 9001-0048
; CURRENT APPLICATION NUMBER: US/09/305,924A
; EARLIER FILING DATE: 1999-05-05
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PRT
; ORGANISM: GNRH
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)
; OTHER INFORMATION: Xaa is pyroglutamic acid
US-09-305-924-5

Query Match 99.0%; Score 100; DB 17; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPGRRPPPC 16
Db 2 HWSYGLRPGRRPPPC 16

RESULT 4
US-09-306-689-5
; Sequence 5, Application US/09306689B
; GENERAL INFORMATION:
; APPLICANT: Robbins, Sarah C.
; TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
; FILE REFERENCE: 9001-0047

;; CURRENT APPLICATION NUMBER: US/09/306,689B
;; CURRENT FILING DATE: 1999-05-06
;; EARLIER APPLICATION NUMBER: US 60/088,024
;; EARLIER FILING DATE: 1998-06-04
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO: 5
;; LENGTH: 16
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; NAME/KEY: MOD_RES
;; LOCATION: (1)
;; OTHER INFORMATION: Xaa is pyroglutamic acid
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Sequence
;; OTHER INFORMATION: defines GNRH analogue
;; US-09-306-689-5

Query Match 99.0%; Score 100; DB 17; Length 16;
Best Local Similarity 100.0%; Pred. No. 1,7e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPGRRPPPC 16
|||||
DB 2 HWSYGLRPGRRPPPC 16

RESULT 5
PCT-US95-01225-6
;; Sequence 6, Application PC/TUS9501225
;; GENERAL INFORMATION:
;; APPLICANT: Grimes, Stephen
;; APPLICANT: Scibienski, Robert
;; TITLE OF INVENTION: Immunogens Against Gonadotropin
;; TITLE OF INVENTION: Releasing Hormone
;; NUMBER OF SEQUENCES: 11
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Dimitrios T. Drivas, Esq.
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10036-2787
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/01225
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Drivas Esq., Dimitrios T.
;; REGISTRATION NUMBER: 32,218
;; REFERENCE/DOCKET NUMBER: 1102865-300
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-819-8286
;; TELEFAX: 212-354-8113
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 17 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: YES
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 1
;; OTHER INFORMATION: /label="pglu
;; OTHER INFORMATION: /note="pyroglutamic acid (5-oxoprolinone)"

;; FEATURE:
;; NAME/KEY: Region
;; LOCATION: 1..10
;; OTHER INFORMATION: /note="immunomimic"
;; FEATURE:
;; NAME/KEY: Region
;; LOCATION: 11..17
;; OTHER INFORMATION: /note="spacer"
;; PCT-US95-01225-6

Query Match 82.7%; Score 83.5; DB 1; Length 17;
Best Local Similarity 87.5%; Pred. No. 0.0017;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 2 HWSYGLRPG-RPPPC 16
|||||
DB 2 HWSYGLRPGSSPPPC 17

RESULT 6
US-08-478-546B-6
;; Sequence 6, Application US/08478546B
;; GENERAL INFORMATION:
;; APPLICANT: Grimes, Stephen
;; APPLICANT: Scibienski, Robert
;; TITLE OF INVENTION: Methods for the Treatment of Hormone-Dependent
;; TITLE OF INVENTION: Tumors with Immunogens Against Gonadotropin Releasing Hormo
;; NUMBER OF SEQUENCES: 11
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Dimitrios T. Drivas, Esq.
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10036-2787
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/478,546B
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/188,223
;; FILING DATE: 27-JAN-1994
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Drivas Esq., Dimitrios T.
;; REGISTRATION NUMBER: 32,218
;; REFERENCE/DOCKET NUMBER: 1102865-300
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-819-8286
;; TELEFAX: 212-354-8113
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 17 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FEATURE:
;; NAME/KEY: Modified site
;; LOCATION: 1
;; OTHER INFORMATION: /note="Xaa
;; OTHER INFORMATION: /note="pyroglutamic acid"
;; FEATURE:
;; NAME/KEY: Region
;; LOCATION: 1..10
;; OTHER INFORMATION: /note="immunomimic"
;; FEATURE:
;; NAME/KEY: Region

LOCATION: 11..17
OTHER INFORMATION: /note="spacer"
US-08-478-546B-6

Query Match 82.7%; Score 83.5; DB 8; Length 17;
Best Local Similarity 87.5%; Pred. No. 0.0017;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 2 HWSYGLRPG-RPPPC 16
Db 2 HWSYGLRPGSSPPPC 17

RESULT 7
US-09-305-924-4
Sequence 4, Application US/09305924A
GENERAL INFORMATION:
APPLICANT: Jack G. Manns
APPLICANT: Stephen D. Acres
APPLICANT: Richard Harland
TITLE OF INVENTION: METHODS OF RAISING ANIMALS FOR MEAT PRODUCTION
FILE REFERENCE: 9001-0048
CURRENT APPLICATION NUMBER: US/09/305,924A
CURRENT FILING DATE: 1999-05-05
EARLIER APPLICATION NUMBER: US 60/084,217
EARLIER FILING DATE: 1998-05-05
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 17
TYPE: PRT
ORGANISM: GNRH
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (1)
OTHER INFORMATION: Xaa is pyroglutamic acid
US-09-305-924-4

Query Match 82.7%; Score 83.5; DB 17; Length 17;
Best Local Similarity 87.5%; Pred. No. 0.0017;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 2 HWSYGLRPG-RPPPC 16
Db 2 HWSYGLRPGSSPPPC 17

RESULT 8
US-09-306-689-4
Sequence 4, Application US/09306689B
GENERAL INFORMATION:
APPLICANT: Robbins, Sarah C.
TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
TITLE OF INVENTION: METHODS FOR SUPPRESSING REPRODUCTIVE BEHAVIOR IN
FILE REFERENCE: 9001-0047
CURRENT APPLICATION NUMBER: US/09/306,689B
CURRENT FILING DATE: 1999-05-06
EARLIER APPLICATION NUMBER: US 60/088,024
EARLIER FILING DATE: 1998-06-04
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 17
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (1)
OTHER INFORMATION: Xaa is pyroglutamic acid
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Sequence

OTHER INFORMATION: defines GNRH analogue
US-09-306-689-4

Query Match 82.7%; Score 83.5; DB 17; Length 17;
Best Local Similarity 87.5%; Pred. No. 0.0017;
Matches 14; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 2 HWSYGLRPG-RPPPC 16
Db 2 HWSYGLRPGSSPPPC 17

RESULT 9
US-09-341-590-103
Sequence 103, Application US/09341590
GENERAL INFORMATION:
APPLICANT: Larsen, Bjarne Due
TITLE OF INVENTION: PHARMACOLOGICALLY ACTIVE PEPTIDE CONJUGATES HAVING A
TITLE OF INVENTION: REDUCED TENDENCY TOWARDS ENZYMATIC HYDROLYSIS
FILE REFERENCE: PPT-20479-US
CURRENT APPLICATION NUMBER: US/09/341,590
CURRENT FILING DATE: 1999-07-03
PRIOR APPLICATION NUMBER: DK 0317/98
PRIOR FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 121
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 103
LENGTH: 16
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: GNRH-Lys6
US-09-341-590-103

Query Match 59.4%; Score 60; DB 17; Length 16;
Best Local Similarity 90.0%; Pred. No. 0.98;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG-R 11
Db 2 HWSYGLRPG-R 11

RESULT 10
US-09-341-590-104
Sequence 104, Application US/09341590
GENERAL INFORMATION:
APPLICANT: Larsen, Bjarne Due
TITLE OF INVENTION: PHARMACOLOGICALLY ACTIVE PEPTIDE CONJUGATES HAVING A
TITLE OF INVENTION: REDUCED TENDENCY TOWARDS ENZYMATIC HYDROLYSIS
FILE REFERENCE: PPT-20479-US
CURRENT APPLICATION NUMBER: US/09/341,590
CURRENT FILING DATE: 1999-07-03
PRIOR APPLICATION NUMBER: DK 0317/98
PRIOR FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 121
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 104
LENGTH: 16
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: GNRH-(Lys-Gly)3
US-09-341-590-104

Query Match 59.4%; Score 60; DB 17; Length 16;
Best Local Similarity 90.0%; Pred. No. 0.98;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPG-R 11

Db 2 HWSYGLRPGK 11

RESULT 11
PCT-US00-22121-26
; Sequence 26, Application PC/TUS0022121
; GENERAL INFORMATION:
; APPLICANT: Proteinix Company
; TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: IGN-2004WO
; CURRENT APPLICATION NUMBER: PCT/US00/22121
; CURRENT FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 09/026,276
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 09/374,721
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: polypeptide
PCT-US00-22121-26

Query Match 58.4%; Score 59; DB 1; Length 20;
Best Local Similarity 90.0%; Pred. No. 1.6;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 HWSYGLRPGK 11
Db 2 HWSYGLRPGQ 11

RESULT 12
PCT-US00-22121-30
; Sequence 30, Application PC/TUS0022121
; GENERAL INFORMATION:
; APPLICANT: Proteinix Company
; TITLE OF INVENTION: HEAT SHOCK FUSION-BASED VACCINE SYSTEM
; FILE REFERENCE: IGN-2004WO
; CURRENT APPLICATION NUMBER: PCT/US00/22121
; CURRENT FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 09/026,276
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: US 09/374,721
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: polypeptide
PCT-US00-22121-30

Query Match 58.4%; Score 59; DB 1; Length 20;
Best Local Similarity 90.0%; Pred. No. 1.6;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 HWSYGLRPGK 11
Db 2 HWSYGLRPGQ 11
RESULT 13

US-07-761-849-10
; Sequence 10, Application US/07761849
; GENERAL INFORMATION:
; APPLICANT: Meloen, Robert H.
; TITLE OF INVENTION: WENSING, CORNELIUS J. G.
; TITLE OF INVENTION: PEPTIDE, IMMUNOGENIC COMPOSITION AND VACCINE
; TITLE OF INVENTION: OR MEDICINAL PREPARATION: A METHOD OF
; TITLE OF INVENTION: IMMUNISING A MAMMAL AGAINST LHRH, AND A METHOD
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/761,849
; FILING DATE: 19910917
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Moran, Thomas F.
; REGISTRATION NUMBER: 16,579
; REFERENCE/DOCKET NUMBER: 2805/40057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-07-761-849-10

Query Match 58.4%; Score 59; DB 3; Length 20;
Best Local Similarity 90.0%; Pred. No. 1.6;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 HWSYGLRPGK 11
Db 2 HWSYGLRPGQ 11

RESULT 14
US-08-476-013-10
; Sequence 10, Application US/08476013
; GENERAL INFORMATION:
; APPLICANT: Meloen, Robert H.
; APPLICANT: Wensing, Cornelius J. G.
; TITLE OF INVENTION: PEPTIDE, IMMUNOGENIC COMPOSITION AND VACCINE
; TITLE OF INVENTION: OR MEDICINAL PREPARATION: A METHOD OF
; TITLE OF INVENTION: IMMUNISING A MAMMAL AGAINST LHRH, AND A METHOD
; TITLE OF INVENTION: OF IMPROVING THE MEAT QUALITY OF PIGS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,013
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Katz, Robert D.
REGISTRATION NUMBER: 30,141
REFERENCE/DOCKET NUMBER: 40057-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-476-013-10

Query Match 58.4%; Score 59; DB 8; Length 20;
Best Local Similarity 90.0%; Pred. No. 1.6;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPGK 11
Db 2 HWSYGLRPGQ 11

RESULT 15

US-08-476-013-12

Sequence 12, Application US/08476013A

GENERAL INFORMATION:

APPLICANT: Meisner, Robert H

APPLICANT: Mensing, Cornelius J. G.

TITLE OF INVENTION: A PEPTIDE, IMMUNOGENIC COMPOSITIONS, VACCINES, AND

TITLE OF INVENTION: MEDICINAL PREPARATIONS INCLUDING SAME; AND METHODS

TITLE OF INVENTION: FOR IMMUNIZING A MAMMAL AGAINST LHRH, AND IMPROVING

TITLE OF INVENTION: THE MEAT QUALITY OF PIGS

FILE REFERENCE: 2183-403605

CURRENT APPLICATION NUMBER: US/08/476,013A

CURRENT FILING DATE: 1995-06-07

NUMBER OF SEQ ID NOS: 17

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 12

LENGTH: 20

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: at least two

OTHER INFORMATION: LHRH peptide sequences in tandem

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)

OTHER INFORMATION: Residue may be Pyroglutamic Acid or Gln having

OTHER INFORMATION: attached thereto a tail including one or more

OTHER INFORMATION: additional amino acids

FEATURE:

NAME/KEY: VARIANT

LOCATION: (3)

OTHER INFORMATION: Trp may be formylated

FEATURE:

NAME/KEY: VARIANT

LOCATION: (13)

OTHER INFORMATION: TRP may be formylated

FEATURE:

NAME/KEY: VARIANT

LOCATION: (10) (19)

OTHER INFORMATION: Sequence may repeat creating more than two LHRH

OTHER INFORMATION: sequences in tandem

FEATURE:

NAME/KEY: VARIANT
LOCATION: (10) (11)
OTHER INFORMATION: Gly residue may be joined to Gln residue by a
OTHER INFORMATION: direct bond or by a spacer made of one or more
OTHER INFORMATION: amino acids, a hydrocarbon chain or other groups
FEATURE:
NAME/KEY: VARIANT
LOCATION: (20)
OTHER INFORMATION: Residue may be amidated or may have attached
OTHER INFORMATION: thereto a tail of one or more additional amino
OTHER INFORMATION: acids
US-08-476-013-12

Query Match 58.4%; Score 59; DB 8; Length 20;
Best Local Similarity 90.0%; Pred. No. 1.6;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 HWSYGLRPGK 11
Db 2 HWSYGLRPGQ 11

Search completed: March 2, 2001, 10:59:03
Job time: 373 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2001, 10:53:46 ; Search time 66.16 Seconds
(without alignments)
8.269 Million cell updates/sec

Title: US-09-306-689-5

Perfect score: 101

Sequence: 1 XHMSYGLRPPPPPC 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /SIDSL/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	99.0	16	16	R78285
2	100	99.0	16	21	Y58141
3	83.5	82.7	17	16	R78283
4	83.5	82.7	17	21	Y58140
5	63	62.4	15	15	R44955
6	63	62.4	15	15	R44956
7	59	58.4	20	19	W47438
8	59	58.4	20	20	Y31174
9	59	58.4	20	20	Y31178
10	59	58.4	21	11	R07324
11	59	58.4	30	11	R07323
12	59	58.4	40	20	Y31183

13	59	58.4	41	20	Y31182
14	58	57.4	9	20	W94891
15	58	57.4	10	2	P10097
16	58	57.4	10	2	P10411
17	58	57.4	10	2	P10416
18	58	57.4	10	6	P50222
19	58	57.4	10	7	P60127
20	58	57.4	10	7	P61403
21	58	57.4	10	7	P60576
22	58	57.4	10	8	P70922
23	58	57.4	10	10	P90630
24	58	57.4	10	12	R15713
25	58	57.4	10	13	R26819
26	58	57.4	10	15	R62889
27	58	57.4	10	16	R91197
28	58	57.4	10	16	R86845
29	58	57.4	10	16	R75152
30	58	57.4	10	17	W65201
31	58	57.4	10	17	W65203
32	58	57.4	10	18	W45842
33	58	57.4	10	18	W22390
34	58	57.4	10	18	W16390
35	58	57.4	10	18	W04612
36	58	57.4	10	19	W79566
37	58	57.4	10	19	W76381
38	58	57.4	10	19	W76373
39	58	57.4	10	19	W61541
40	58	57.4	10	20	Y50229
41	58	57.4	10	20	Y31176
42	58	57.4	10	20	Y31180
43	58	57.4	10	20	Y31057
44	58	57.4	10	20	Y03864
45	58	57.4	10	21	Y55061

ALIGNMENTS

RESULT 1	
R78285	
ID	R78285 standard; peptide; 16 AA.
XX	
AC	R78285;
XX	
DT	13-MAR-1996 (first entry)
XX	
DE	GnRH immunomimetic and spacer (GnRH(1-10)-Arg10).
XX	
KW	Immunomimetic; gonadotropin releasing hormone; GnRH; LHRH; vaccine;
KW	luteinizing hormone releasing hormone; GnRH; immunomimetic; uterine;
KW	diphtheria toxin; DT; gynaecological; endometriosis; uterine fibroids;
KW	gonadal steroid hormone associated dependent disease; gonadotropin;
KW	immunological contraception; mammary; breast; cancer; prostate;
KW	benign prostatic hypertrophy.
XX	
OS	Chimeric - Homo sapiens.
OS	Chimeric - Synthetic.
XX	
FT	Key
FT	Modified-site
FT	Location/Qualifiers
FT	1
FT	/label= OTHER
FT	/note= "pyroglutamic acid"
FT	1..10
FT	/note= "GnRH"
FT	11..16
FT	/note= "spacer"
XX	
PN	W09520600-A1.
XX	
PD	03-AUG-1995.
XX	
PF	26-JAN-1995; 95WO-US01225.
XX	

Ubiqutin fusion p
LHRH peptide fragm
Sequence of lutein
luteinizing Hormon
luteinizing Hormon
Gonadotropin rele
Gonadoliberin anta
Gonadotropin relea
Novel decapeptide
luteinizing hormon
Sequence of lutein
Peptide #1 with ho
LH releasing hormo
LHRH hapten for at
LHRH peptide. Syn
Gonadotropin relea
Gonadotropin relea
luteinizing hormon
Luteinizing hormon
Gonadotropin relea
Gonadotropin relea
Luteinizing hormon
Novel decapeptide
GnRH-1 polypeptide
Rat modified GnRH
Rat GnRH peptide.
Peptide hormone Gn
Neutrophil-activat
Ubiqutin fusion p
Ubiqutin fusion p
Non-crosslinked pr
Amino acid sequenc
luteinizing hormon

CC The sequence is that of an insert contg. the decapeptide gonadotropin
 CC releasing hormone (GnRH) located in the major subunit of P-fimbriae
 CC with serotype F11 at a position corresponding to a position in
 CC hypervariable region 4 of the wild type major subunit. This forms
 CC part of an immunogenic carrier system capable of eliciting an immune
 CC response against GnRH.

XX Sequence 15 AA;

Query Match 62.4%; Score 63; DB 15; Length 15;

Best Local Similarity 90.9%; Pred. No. 0.033; 1; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPRG 12
 |||||
 Db 4 hwsyglrpgdp 14

RESULT 6

R44956 R44956 standard; Protein; 16 AA.

XX R44956;

XX 07-JUL-1994 (first entry)

XX E. coli fimbriae HR4 GnRH insert linker.

XX Gonadotropin releasing hormone; immunogenic carrier system;

XX Immune response; immunisation; vaccine; contraceptive agent;

XX treatment; sexual hyperactivity; cancer.

XX Synthetic.

XX Key

XX Peptide

XX Location/Qualifiers

XX 2..11

XX /note= "decapeptide GnRH"

XX EP578293-A.

XX 12-JAN-1994.

XX 15-JUN-1993; 93EP-0201712.

XX 18-JUN-1992; 92EP-0201775.

XX (ALKU) AKZO NV.

XX Gielen JT, Hoekstra WPM, Van Der Zee A, Van Die IM;

XX WPI; 1994-010165/02.

XX N-PSDB; Q54395.

XX Immunogenic carrier system comprising Escherichia coli P-fimbriae

XX - used to elicit immune response against gonadotropin releasing

XX hormone

XX Example 1; Fig 2; 34pp; English.

XX The sequence is that of an insert contg. the decapeptide gonadotropin

XX releasing hormone (GnRH) located in the major subunit of P-fimbriae

XX with serotype F11 at a position corresponding to a position in

XX hypervariable region 4 of the wild type major subunit. This forms

XX part of an immunogenic carrier system capable of eliciting an immune

XX response against GnRH.

XX Sequence 16 AA;

Query Match 62.4%; Score 63; DB 15; Length 16;

Best Local Similarity 90.9%; Pred. No. 0.035; 1; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HWSYGLRPRG 12
 |||||
 Db 5 hwsyglrpgdp 15

RESULT 7

W47438 W47438 standard; peptide; 20 AA.

XX W47438;

XX 05-JUN-1998 (first entry)

XX Antigenic peptide.

XX Vaccine; antigen.

XX Synthetic.

XX WO9749425-A1.

XX 31-DEC-1997.

XX 24-JUN-1997; 97WO-NL00354.

XX 25-JUN-1996; 96EP-0201766.

XX (DAVE-) DANISH VETERINARY INST ANIMAL VIRUS RES.

XX (DIER-) STICHTING INST DIERHOUDERTJ EN DIERGEZONDHEID.

XX Beekman NJCM, Dalsgaard K, Melsen RH, Schaaper WMM;

XX WPI; 1998-076912/07.

XX Vaccines comprising antigen bound to carrier by an in vivo labile

XX bond - especially synthetic peptide linked to fatty acid via

XX thioester or di-sulphide, provide greater immune response for weakly

XX immunogenic antigens

XX Claim 9; Page 28; 36pp; English.

XX A novel vaccine comprises an antigen (Ag), e.g. the present

XX peptide, and carrier connected by a bond that is labile and

XX dissociates under certain physiological conditions.

XX The vaccine, which allows dissociation of the Ag from the carrier

XX molecule, can be used to elicit better immune responses against

XX poorly immunogenic Ag then those Ag which contain a stable link to

XX the carrier molecule. Ag dissociate from the carrier in vivo,

XX resulting in better immune response for Ag that are normally only

XX weakly immunogenic. The vaccine also improves targeting to, and

XX presentation by Ag-presenting cells.

XX Sequence 20 AA;

Query Match 58.4%; Score 59; DB 19; Length 20;

Best Local Similarity 90.0%; Pred. No. 0.14; 1; Indels 0; Gaps 0;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPRG 11
 |||||
 Db 2 hwsyglrpgdp 11

RESULT 8

Y31174 Y31174 standard; peptide; 20 AA.

XX Y31174;

XX 28-OCT-1999 (first entry)

DE Ubiquitin fusion protein GnRH dimer for C-terminal extension.
 XX
 KW Ubiquitin; immunocastration; fusion protein; heat shock protein; epitope;
 KW immune response stimulation; vaccine; T cell; viral; infection; cancer;
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;
 KW pulmonary infection; respiratory infection; scaffold; anti-self; pig;
 KW steridogenesis; gamete maturation; prostate; breast; castration; TNF;
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;
 KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;
 KW fertility; sperm protein; growth rate; antibody; detection; GnRH.
 XX
 OS Synthetic.
 XX
 PN WO9942472-A1.
 XX
 PD 26-AUG-1999.
 XX
 PF 26-JAN-1999; 99WO-US01588.
 XX
 PR 19-FEB-1998; 98US-0026276.
 XX
 PA (IGEN-) IGEN INT INC.
 XX
 PI Kenten JH, Lohas GL, Pilon AL, Roberts SF, Tramontano A;
 XX
 DR WPI: 1999-518582/43.
 XX
 PT Epitope-containing fusion proteins used to generate a highly
 XX specific immune responses
 XX
 PS Example 3; Page 40; 67pp; English.
 XX
 CC This invention describes a novel fusion protein, comprising a heat shock
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner
 CC which is useful for the stimulation of a highly specific immune response
 CC when administered to an animal. The protein of the invention may be
 CC post-translationally modified (e.g. by the addition of fatty acids to
 CC enhance immunogenicity). The fusion proteins of the invention can be
 CC used as vaccines to induce an immune response. When a T cell epitope is
 CC attached, they can be used for control of viral infections, bacterial
 CC infections, parasitic infection and cancer. The fusion proteins can be
 CC used in pharmaceutical compositions for the treatment of gastrointestinal
 CC diseases, pulmonary infections, respiratory infections, and HIV
 CC infections. The use of ubiquitin as a scaffold is also useful for the
 CC presentation and stimulation of anti-self immune responses, e.g.
 CC generation of anti-gonadotropin releasing hormone antibodies which result
 CC in the suppression of luteinizing hormone and follicle stimulating
 CC hormone. This indirectly suppresses steridogenesis and gamete maturation
 CC in males and females. This type of anti-self response in humans is useful
 CC in the treatment of prostate cancer and breast cancer. In livestock, the
 CC ability to stimulate an anti-self response provides a simple alternative
 CC to physical castration. Immunocastration of pigs is a better alternative
 CC to physical castration, as it does not result in any of the detrimental
 CC side effects associated with physical castration. Other examples of
 CC diseases and conditions treated with self proteins fused with ubiquitin
 CC are TNF and its epitopes to modulate septic shock, arthritis,
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig
 CC epsilon heavy chain for the control of allergic reactions; chorionic
 CC gonadotropin for fertility control; and sperm proteins for fertility
 CC control. A further use of the fusion proteins is as part of a vaccine to
 CC enhance growth rate and thereby the final weight of the livestock prior
 CC to shipment to market. In addition, the fusion proteins of the invention
 CC can be used to detect and identify antibodies from experimental samples.
 CC This sequence represents a GnRH dimer used in the construction of a
 CC ubiquitin fusion protein described in the method of the invention.
 XX
 SQ Sequence 20 AA;

Query Match 58.4%; Score 59; DB 20; Length 20;
 Best Local Similarity 90.0%; Pred. No. 0.14;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 HWSYGLRPCR 11
 Db 2 hwsyglrpgq 11
 |||||
 RESULT 9
 Y31178
 ID Y31178 standard; peptide: 20 AA.
 XX
 AC Y31178;
 XX
 DT 28-OCT-1999 (first entry)
 XX
 DE Ubiquitin fusion protein GnRH mixed dimer 1.
 XX
 KW Ubiquitin; immunocastration; fusion protein; heat shock protein; epitope;
 KW immune response stimulation; vaccine; T cell; viral; infection; cancer;
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;
 KW pulmonary infection; respiratory infection; scaffold; anti-self; pig;
 KW steridogenesis; gamete maturation; prostate; breast; castration; TNF;
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;
 KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;
 KW fertility; sperm protein; growth rate; antibody; detection; GnRH.
 XX
 OS Synthetic.
 XX
 PN WO9942472-A1.
 XX
 PD 26-AUG-1999.
 XX
 PF 26-JAN-1999; 99WO-US01588.
 XX
 PR 19-FEB-1998; 98US-0026276.
 XX
 PA (IGEN-) IGEN INT INC.
 XX
 PI Kenten JH, Lohas GL, Pilon AL, Roberts SF, Tramontano A;
 XX
 DR WPI: 1999-518582/43.
 XX
 PT Epitope-containing fusion proteins used to generate a highly
 XX specific immune responses
 XX
 PS Example 3; Page 41; 67pp; English.
 XX
 CC This invention describes a novel fusion protein, comprising a heat shock
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner
 CC which is useful for the stimulation of a highly specific immune response
 CC when administered to an animal. The protein of the invention may be
 CC post-translationally modified (e.g. by the addition of fatty acids to
 CC enhance immunogenicity). The fusion proteins of the invention can be
 CC used as vaccines to induce an immune response. When a T cell epitope is
 CC attached, they can be used for control of viral infections, bacterial
 CC infections, parasitic infection and cancer. The fusion proteins can be
 CC used in pharmaceutical compositions for the treatment of gastrointestinal
 CC diseases, pulmonary infections, respiratory infections, and HIV
 CC infections. The use of ubiquitin as a scaffold is also useful for the
 CC presentation and stimulation of anti-self immune responses, e.g.
 CC generation of anti-gonadotropin releasing hormone antibodies which result
 CC in the suppression of luteinizing hormone and follicle stimulating
 CC hormone. This indirectly suppresses steridogenesis and gamete maturation
 CC in males and females. This type of anti-self response in humans is useful
 CC in the treatment of prostate cancer and breast cancer. In livestock, the
 CC ability to stimulate an anti-self response provides a simple alternative
 CC to physical castration. Immunocastration of pigs is a better alternative
 CC to physical castration, as it does not result in any of the detrimental
 CC side effects associated with physical castration. Other examples of
 CC diseases and conditions treated with self proteins fused with ubiquitin
 CC are TNF and its epitopes to modulate septic shock, arthritis,
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig
 CC epsilon heavy chain for the control of allergic reactions; chorionic
 CC gonadotropin for fertility control; and sperm proteins for fertility
 CC control. A further use of the fusion proteins is as part of a vaccine to

CC enhance growth rate and thereby the final weight of the livestock prior
 CC to shipment to market. In addition, the fusion proteins of the invention
 CC can be used to detect and identify antibodies from experimental samples.
 CC This sequence represents a GnRH mixed dimer used in the construction of
 CC a ubiquitin fusion protein described in the method of the invention.

XX Sequence 20 AA;

Query Match 58.4%; Score 59; DB 20; Length 20;
 Best Local Similarity 90.0%; Pred. No. 0.14;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPR 11
 |||||
 Db 2 hwsyglrpg 11

RESULT 10
 R07324 R07324 standard; protein; 21 AA.

AC R07324;

DT 29-JAN-1991 (first entry)

DE Lutetinsing hormone releasing hormone derived peptide.

XX LHRH; vaccine; meat; pigs; cancer; sterilisation.

XX Synthetic.

OS Key Location/Qualifiers

FT Modified-site 1 /label=OTHER

FT Modified-site 3 /note="OTHER-pyroglyutamic acid"

FT Modified-site 13 /label=OTHER

FT Modified-site 13 /note="OTHER- N-formyl-Trp (optional)"

FT Misc-difference 21 /label=OTHER

FT /note="OTHER- N-formyl-Trp (optional)"

XX WO9011298-A.

XX 04-OCT-1990.

XX 22-MAR-1990; 90WO-NL00037.

XX 23-MAR-1989; 89NL-0000726.

XX (DIER-) STICHT CENT DIERGEN.

XX Meloen RH, Wensing CUG;

XX WPI; 1990-320228/42.

PT Peptide for vaccinating mammals against LHRH - comprises at least
 PT two lutetinsing hormone releasing hormone sequences in tandem

PS Claim 4; Page 10; 15pp; English.

CC The peptide comprises at least 2 LHRH sequences in tandem. The
 CC peptide can be used to vaccinate mammals (e.g. pigs) against LHRH.
 CC Such vaccination is used in human medicine for the treatment of
 CC prostate cancer and breast cancer and some forms of hypophyseal
 CC carcinoma. Other applications include sterilisation of domestic
 CC animals and treatment of aggression in dogs. A major use of the
 CC vaccination is to improve meat quality in pigs by avoiding "boar
 CC odour" associated with the meat of sexually mature pigs.

CC See also R07323.
 XX Sequence 21 AA;

Query Match 58.4%; Score 59; DB 11; Length 21;
 Best Local Similarity 90.0%; Pred. No. 0.14;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPR 11
 |||||
 Db 2 hwsyglrpg 11

RESULT 11
 R07323 R07323 standard; peptide; 30 AA.

AC R07323;

DT 29-JAN-1991 (first entry)

DE Lutetinsing hormone releasing hormone derived peptide.

XX LHRH; vaccine; meat; pigs; cancer; sterilisation.

XX Synthetic.

OS Key Location/Qualifiers

FT Misc-difference 1 /label=OTHER

FT /note="OTHER-pyroglyutamic acid or Gln having at
 FT least one additional AA attached."

FT Modified-site 3 /label=OTHER

FT Modified-site 13 /note="OTHER- N-formyl-Trp (optional)"

FT Modified-site 13 /label=OTHER

FT Modified-site 10..19 /note="OTHER- N-formyl-Trp (optional)"

FT Region /label=repeat

FT Misc-difference 30 /note="repeat must occur at least once"

FT /label=OTHER

FT /note="OTHER-Gly-NH2 or Gly having at
 FT least one additional AA attached"

XX WO9011298-A.

XX 04-OCT-1990.

XX 22-MAR-1990; 90WO-NL00037.

XX 23-MAR-1989; 89NL-0000726.

XX (DIER-) STICHT CENT DIERGEN.

XX Meloen RH, Wensing CUG;

XX WPI; 1990-320228/42.

PT Peptide for vaccinating mammals against LHRH - comprises at least
 PT two lutetinsing hormone releasing hormone sequences in tandem

PS Claim 2; Page 10; 15pp; English.

CC The peptide comprises at least 2 LHRH sequences in tandem. There
 CC may be a spacer gp. between Gly(20) and Gln(21). The peptide can
 CC be used to vaccinate mammals (e.g. pigs) against LHRH. Such
 CC vaccination is used in human medicine for the treatment of prostate
 CC cancer and breast cancer and some forms of hypophyseal carcinoma.
 CC Other applications include sterilisation of domestic animals and
 CC treatment of aggression in dogs. A major use of the vaccination is

CC / to improve meat quality in pigs by avoiding "boar odour" associated
 CC with the meat of sexually mature pigs.
 CC See also R07323.

XX Sequence 30 AA;

Query Match 58.4%; Score 59; DB 11; Length 30;
 Best Local Similarity 90.0%; Pred. No. 0.19;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 HMSYGLRPCR 11
 |||||
 Db 2 hwsyglrpgq 11

RESULT 12
 Y31183
 ID Y31183 standard; peptide: 40 AA.

AC Y31183;
 XX
 DT 28-OCT-1999 (first entry)

DE Ubiquitin fusion protein GnRH fragment 2.

XX
 KW Ubiquitin; immunocastration; fusion protein; heat shock protein; epitope;
 KW immune response stimulation; vaccine; T cell; viral; infection; cancer;
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;
 KW pulmonary infection; respiratory infection; scaffold; anti-self; pig;
 KW steriodogenesis; gamete maturation; prostate; breast; castration; TNF;
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;
 KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;
 KW fertility; sperm protein; growth rate; antibody; detection; GnRH.

XX Unidentified.

OS

PN W09942472-A1.

PD 26-AUG-1999.

PF 26-JAN-1999; 99WO-US01588.

PR 19-FEB-1998; 98US-0026276.

PA (IGEN-) IGEN INT INC.

PI Kenten JH, Lohnas GL, Pilon AL, Roberts SF, Tramontano A;
 WPI: 1999-518582/43.

PT Epitope-containing fusion proteins used to generate a highly
 specific immune responses

PS Claim 83; Page 43; 67pp; English.

XX This invention describes a novel fusion protein, comprising a heat shock
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner
 CC which is useful for the stimulation of a highly specific immune response
 CC when administered to an animal. The protein of the invention may be
 CC post-translationally modified (e.g. by the addition of fatty acids to
 CC enhance immunogenicity). The fusion proteins of the invention can be
 CC used as vaccines to induce an immune response. When a T cell epitope is
 CC attached, they can be used for control of viral infections, bacterial
 CC infections, parasitic infection and cancer. The fusion proteins can be
 CC used in pharmaceutical compositions for the treatment of gastrointestinal
 CC diseases, pulmonary infections, respiratory infections, and HIV
 CC infections. The use of ubiquitin as a scaffold is also useful for the
 CC presentation and stimulation of anti-self immune responses, e.g.
 CC generation of anti-gonadotropin releasing hormone antibodies which result
 CC in the suppression of luteinizing hormone and follicle stimulating
 CC hormone. This indirectly suppresses steriodogenesis and gamete maturation
 CC in males and females. This type of anti-self response in humans is useful

CC in the treatment of prostate cancer and breast cancer. In livestock, the
 CC ability to stimulate an anti-self response provides a simple alternative
 CC to physical castration. Immunocastration of pigs is a better alternative
 CC to physical castration, as it does not result in any of the detrimental
 CC side effects associated with physical castration. Other examples of
 CC diseases and conditions treated with self proteins fused with ubiquitin
 CC are TNF and its epitopes to modulate septic shock, arthritis,
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig
 CC epsilon heavy chain for the control of allergic reactions; chorionic
 CC gonadotropin for fertility control; and sperm proteins for fertility
 CC control. A further use of the fusion proteins is as part of a vaccine to
 CC enhance growth rate and thereby the final weight of the livestock prior
 CC to shipment to market. In addition, the fusion proteins of the invention
 CC can be used to detect and identify antibodies from experimental samples.
 CC This sequence represents a GnRH fragment used in the construction of
 CC a ubiquitin fusion protein described in the method of the invention.

XX Sequence 40 AA;

Query Match 58.4%; Score 59; DB 20; Length 40;
 Best Local Similarity 90.0%; Pred. No. 0.25;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 HMSYGLRPCR 11
 |||||
 Db 2 hwsyglrpgq 11

RESULT 13
 Y31182
 ID Y31182 standard; peptide: 41 AA.

AC Y31182;
 XX
 DT 28-OCT-1999 (first entry)

DE Ubiquitin fusion protein GnRH fragment.

XX
 KW Ubiquitin; immunocastration; fusion protein; heat shock protein; epitope;
 KW immune response stimulation; vaccine; T cell; viral; infection; cancer;
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;
 KW pulmonary infection; respiratory infection; scaffold; anti-self; pig;
 KW steriodogenesis; gamete maturation; prostate; breast; castration; TNF;
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;
 KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;
 KW fertility; sperm protein; growth rate; antibody; detection; GnRH.

XX Unidentified.

OS

PN W09942472-A1.

PD 26-AUG-1999.

PF 26-JAN-1999; 99WO-US01588.

PR 19-FEB-1998; 98US-0026276.

PI Kenten JH, Lohnas GL, Pilon AL, Roberts SF, Tramontano A;
 WPI: 1999-518582/43.

PT Epitope-containing fusion proteins used to generate a highly
 specific immune responses

PS Claim 81; Page 43; 67pp; English.

XX This invention describes a novel fusion protein, comprising a heat shock
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner
 CC which is useful for the stimulation of a highly specific immune response
 CC when administered to an animal. The protein of the invention may be

CC post-translationally modified (e.g. by the addition of fatty acids to
CC enhance immunogenicity). The fusion proteins of the invention can be
CC used as vaccines to induce an immune response. When a T cell epitope is
CC attached, they can be used for control of viral infections, bacterial
CC infections, parasitic infection and cancer. The fusion proteins can be
CC used in pharmaceutical compositions for the treatment of gastrointestinal
CC diseases, pulmonary infections, respiratory infections, and HIV
CC infections. The use of ubiquitin as a scaffold is also useful for the
CC presentation and stimulation of anti-self immune responses, e.g.
CC generation of anti-gonadotropin releasing hormone antibodies which result
CC in the suppression of luteinizing hormone and follicle stimulating
CC hormone. This indirectly suppresses steroidogenesis and gamete maturation
CC in males and females. This type of anti-self response in humans is useful
CC in the treatment of prostate cancer and breast cancer. In livestock, the
CC ability to stimulate an anti-self response provides a simple alternative
CC to physical castration. Immunocastration of pigs is a better alternative
CC to physical castration, as it does not result in any of the detrimental
CC side effects associated with physical castration. Other examples of
CC diseases and conditions treated with self proteins fused with ubiquitin
CC are TME and its epitopes to modulate septic shock, arthritis,
CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis. Ig
CC epsilon heavy chain for the control of allergic reactions; chorionic
CC gonadotropin for fertility control; and sperm proteins for fertility
CC control. A further use of the fusion proteins is as part of a vaccine to
CC enhance growth rate and thereby the final weight of the livestock prior
CC to shipment to market. In addition, the fusion proteins of the invention
CC can be used to detect and identify antibodies from experimental samples.
CC This sequence represents a gNH fragment used in the construction of
CC a ubiquitin fusion protein described in the method of the invention.

XX Sequence 4.1AA:

Query Match	58.4%;	Score 59;	DB 20;	Length 41;
Best Local Similarity	90.0%;	Pred. No. 0.26;		
Matches 9; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;
OY	2	HWSYGLRPGR	11	.
db	2	hwsyglrpgq	11	

RESULT	14
ID	W94891
AC	W94891 standard; peptide; 9 AA.
DT	11-MAY-1999 (first entry)
DE	LHRH peptide fragment.
KW	LHRH; immune response; luteinising hormone releasing hormone; DT; diphtheria toxin; castrating; oestrus cycling; aggression; breast; sexual activity; organoleptic; livestock; cell growth; malignant; prostate; ovarian; oncofetal; hyperplastic; pregnancy; endometriosis; inflammatory response.
OS	Homo sapiens.
PN	WO9902180-A1.
PD	21-JAN-1999.
PF	09-JUL-1998; 98WO-AU00532.
PR	09-JUL-1997; 97AU-0007768.
PA	*(CSLC-) CSL LTD.
PI	McNamara MK;
DR	WPI: 1999-120511/10.

XX New immunogenic leutenising hormone releasing hormone compositions -
PT comprise LHRH conjugated to diptheria toxoid and adsorbed to an
PT ionic polysaccharide, used to inhibit reproductive function in
PT animals
XX
XX
PS Example 3; Page 30; 41pp: English.
XX
XX
CC The invention relates immunogenic composition for eliciting an immune
CC response to luteinising hormone releasing hormone (LHRH). The composition
CC comprises a LHRH-diptheria toxoid (PT) conjugate adsorbed to an ionic
CC polysaccharide. The LHRH-PT compositions can be used for eliciting an
CC immune response to LHRH, for castrating an animal, for regulating an
CC oestrus cycling in a female animal or for inhibiting characteristics
CC induced by the sexual maturation of an animal, e.g. aggression or sexual
CC activity. They can also be used for achieving production gains in
CC livestock, e.g. reduction or elimination of unwanted orandolepic
CC characteristics from the meat of livestock. They can also be used for
CC inhibiting the growth of cells which are regulated directly or indirectly
CC by LHRH, e.g. malignant breast cells, malignant prostate cells, malignant
CC ovarian cells, malignant oncofoetal cells or hyperplastic cells. They can
CC also be used for down-regulating the libido of an animal. They can also
CC be used for inhibiting pregnancy, prostate enlargement, endometriosis or
CC inflammatory responses. The LHRH compositions induce a more effective
CC immune response against LHRH than the LHRH-carrier-adjutant compositions.
CC The effective immune response against LHRH results in prevention of the
CC release of the hormones LH and FSH from the anterior pituitary. Sequences
XX W94890-93 are peptide derivatives of LHRH.
XX
XX Sequence 9 AA;
XX

```

Query March      57.4%; Score 58; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 2,1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0;
QY 2 HWSTGLRPG 10
    |||||
db 1 hwsyglrpg 9

```

RESULT	15
PI0097	
ID	PI0097 standard; peptide; 10 AA.
XX	
AC	PI0097;
XX	
DT	19-AUG-1992 (first entry)
XX	
DE	Sequence of luteinising hormone (LH-RH, ICSH) liberating hormone
XX	
KW	Gonadotropin; luteinising hormone releasing hormone; LH-RH;
KW	ICSH; prostatic hyperplasia therapy.
XX	
OS	Mammal.
XX	
Key	Location/Qualifiers
FH	Misc-difference 1
FT	/label= Pyr
FT	10
Modified-site	/label= Gly-NH2
FT	
XX	
BE887639-A.	
FN	
XX	
24-AUG-1981.	
XX	
27-AUG-1981;	81BE-0303944.
XX	
22-MAY-1980;	80US-0152241.
XX	
PA	(AMHP) AYERST MCKENNA HARR.
XX	
PI	Auclair C;

XX / WPI: 1981-66067D/37 (66067D).

DR Gonadorelin for treatment of benign prostatic hyperplasia - is
XX the deca-peptide Pyr-His-Tyr-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH2 or
PT luteinizing hormone liberating hormone
XX
XX

PS Claim 1, Page 7, 9pp: French.

XX
CC The inventors claim a compsn. for the redn. or prevention of
CC undesired prostatic growth in males. The compsn. contains a
CC decapeptide (gonadorelin) (P10097) with an appropriate vehicle or
CC support. The compsn. is used for treating e.g. benign prostatic
CC hyperplasia by parenteral admin. in daily doses of 0.035-11.0 (pref.
CC 0.080-2.0) mg/kg. Gonadorelin is the generic name for LH-RH and is
CC described in US3835108. In the example s.c. injection of the
CC compsn. significantly reduced the wt. of seminal vesicles and
CC ventral prostate in rats without affecting the wt. of the
CC testicles.
XX

SO Sequence 10 AA:

Query Match 57.4%; Score 58; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HWSYGLRPG 10
|||
|||
Db 2 hwsyglrpg 10

Search completed: March 2, 2001, 10:53:46
Job time: 911 sec

THIS PAGE BLANK (USPTO)

alignment_scores:
 Quality: 290.00 Length: 49
 Ratio: 5.918 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-306-689-10 x W03944 ..

Align seg 1/1 to: W03944 from: 1 to: 49

```

1 CAGCATTTGAGACTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 17
17 rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34
51 CTAGCGCTGCGTCCGGGTGAGCTCTAGCCAGCATTTGAGAGCTAGCGCTGC 100
|||||
101 GCCCTGGCAGCGGTAGCCAGATTGAGACTACGGCTGCGCGGT 147
|||||
34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49

```

seq_name: /SIDSI/gcgdata/geneseq/geneseq/A11998.DAT.W79567

seq_documentation_block:

ID W79567 standard; Protein: 49 AA.

AC W79567;

DT 24-DEC-1998 (first entry)

DE GnRH-2.

KW Gonadotropin releasing hormone; GnRH; decapeptide; hypothalamus; spacer;
 LH; luteinising hormone; FSH; follicle stimulating hormone; vertebrate;
 pyroGlu; chmera; leukotoxin polypeptide; multimer; vaccine; tumour;
 immunogenic.

OS Synthetic.

PN W09806848-A1.

PD 19-FEB-1998.

PF 08-AUG-1997; 97WO-CA00559.

PR 09-AUG-1996; 96US-0694865.

PA (UYSA-) UNIV SASKATCHEWAN.

PI Manns JG, Potter AA;

DR WPI; 1998-159540/14.

DR N-PSDB; V61529.

Chimeric protein of leukotoxin and gonadotropin releasing hormone
 useful for, e.g. preparation of vaccines for reduction of incidence
 of mammary tumours in mammals

Disclosure: Figure 1B; 118pp; English.

The present sequence represents a recombinantly produced or chemically
 synthesised gonadotropin releasing hormone-2 (GnRH-2) polypeptide, which
 contains four copies of the GnRH decapeptide (V61529) and triplet amino
 acid spacers between each of these sequences. This decapeptide is
 secreted naturally by the hypothalamus which controls release of both the
 luteinising hormone (LH) and the follicle stimulating hormone (FSH) in
 vertebrates. This sequence, as compared to the native peptide, has been
 found to have an N-terminal Gln rather than a pyroGlu residue, and also
 contains substitutions at amino acid residues 15 and 41, whereby His is
 replaced by Asp. This produces an alternating multimeric GnRH sequence
 which is highly immunogenic that can be used in the construction of a
 chimeric protein that comprises a leukotoxin polypeptide, several

CC multimers, and the GnRH sequence. The chimeric protein can be used as a
 CC vaccine to help reduce the incidence of mammary tumours in a mammalian
 CC individual.

XX Sequence 49 AA;

SD

alignment_scores:
 Quality: 290.00 Length: 49
 Ratio: 5.918 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-306-689-10 x W79567 ..

Align seg 1/1 to: W79567 from: 1 to: 49

```

1 CAGCATTTGAGACTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
|||||
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 17
17 rTyrGlyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34
51 CTAGCGCTGCGTCCGGGTGAGCTCTAGCCAGCATTTGAGAGCTAGCGCTGC 100
|||||
101 GCCCTGGCAGCGGTAGCCAGATTGAGACTACGGCTGCGCGGT 147
|||||
34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49

```

seq_name: /SIDSI/gcgdata/geneseq/geneseq/A11998.DAT.W61542

seq_documentation_block:

ID W61542 standard; Protein: 49 AA.

AC W61542;

DT 27-OCT-1998 (first entry)

DE Peptide hormone GnRH-2 decapeptide (4 copies) fragment.

KW GnRH; gonadotropin releasing hormone peptide hormone; leukotoxin;
 immunisation; endogenous molecule; vaccine; ear; immunogen; carrier;
 immune response; hormone receptor; cancerous cell; domestic animal;
 porcine; bovine; luteinizing hormone; follicle stimulating hormone;
 immunocastate.

OS Synthetic.

PN W09834639-A1.

PD 13-AUG-1998.

PF 04-FEB-1998; 98WO-CA00059.

PR 05-FEB-1997; 97US-0036883.

PA (BIOS-) BIOSTAR INC.

PI Acres SD, Harland R, Manns JG;

DR WPI; 1998-446952/38.

DR N-PSDB; V45190.

Immunisation against endogenous molecules by administering vaccine
 to ear - useful to elicit efficient and uniform immune response
 against e.g. gonadotropin releasing hormone to immunocastate pigs
 and cattle

Example 1; Fig 1B; 61pp; English.

This represents the amino acid sequence of the gonadotropin releasing
 hormone (GnRH-2) decapeptide fragment used in the chimeric leukotoxin-
 GnRH polypeptide gene fusions. This is used to exemplify the method of

CC invention of immunisation against endogenous molecules by administering
 CC a vaccine which comprises an immunogen and a carrier to the ear of the
 CC mammal. The method is useful for eliciting an efficient and uniform
 CC immune response to block or suppress the activity of an endogenous
 CC hormone, hormone receptor, agonist or antagonist in a vaccinated subject,
 CC or to elicit an immune response against a targeted endogenous cell type
 CC (e.g. a cancerous or otherwise diseased cell). It is especially useful
 CC to reduce the levels of GnRH in domestic animals, especially in porcine
 CC or bovine species. The use of GnRH immunogens in the vaccine reduces the
 CC levels of luteinizing hormone and follicle stimulating hormone and helps
 CC in immunocastrating the animal. Administration of vaccine compositions to
 CC the ear instead of intramuscular administration into the neck increases
 CC the efficiency of vaccination of mammals against endogenous immunogens,
 CC and may increase uniformity of vaccine presentation since the ear is
 CC relatively uniform from animal to animal.

XX Sequence 49 AA:

alignment_scores:

Quality: 290.00 Length: 49
 Ratio: 5.918 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-306-689-10 x W61542 ..

Align seg 1/1 to: W61542 from: 1 to: 49

1 CAGCATTTGAGCTACGCGCTCGCCCTGGCAGCGGTCTTCAGATTGGAG 50
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1 GlnH1StrpSerTyrGlyLeuArgProGlySerGlnAspTrpSe 17
 51 CTACGGCCTGCGTCCGGGTGCTCTAGCCAGCATTTGAGCTCGCCCTGC 100
 ||||||||||||||||||||||||||||||||||||||||||||||||
 17 rTyrGlyLeuArgProGlyGlySerSerGlnH1StrpSerTyrGlyLeuA 34
 101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGCGCTCGCCGGGT 147
 ||||||||||||||||||||||||||||||||||||||||||||||||
 34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49
 seq_name: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:Y58363
 seq_documentation_block:
 ID Y58363 standard; Protein: 49 AA.
 AC Y58363;
 XX
 DT 27-MAR-2000 (first entry)
 XX
 DE Four-copy gonadotropin-releasing hormone (GnRH) multimer.
 XX
 KW GnRH multimer; gonadotropin-releasing hormone; immunosterilisation;
 KM immunoc contraception; vaccine; feline; canine; equine; cervine; ds
 XX
 OS Mammalia.
 XX Synthetic.
 XX
 PN WO962545-A2.
 PD 09-DEC-1999.
 XX
 PF 28-MAY-1999; 99WO-CA00493.
 XX
 PR 04-JUN-1998; 98US-0088024.
 PR 06-MAY-1999; 99US-0306689.
 XX
 PA (BIOS-) BIOSTAR INC.
 XX
 PI Robbins SC;
 XX
 WP1; 2000-086857/07.
 DR N-PSDB; 255702.

XX
 PT Hormone immunogens, analogues or antibodies used to manufacture
 PT vaccines for suppression of reproductive behavior and fertility in
 PT vertebrates -
 XX
 PS Claim 5; Fig 5B; 88P; English.

CC This sequence represents a four-copy gonadotropin-releasing
 CC hormone (GnRH) multimer, where the second and fourth GnRH sequence
 CC have a His to Asp substitution at position 2 of the GnRH sequence.
 CC The invention relates to GnRH immunogens, analogues or antibodies
 CC that cross-react with endogenous GnRH of a vertebrate. A specifically
 CC claimed immunogenic fusion protein (Y58361) comprises, in the N to
 CC C-terminal direction, a synthetic peptide sequence (Y58364), an eight
 CC copy GnRH multimer (composed of two copies of the 4xGnRH multimer
 CC sequence of Y58363), the LRT protein (which functions as a carrier
 CC protein), and a second eight copy GnRH multimer. The fusion protein may
 CC be used in a vaccine composition for prepubertal administration to a
 CC vertebrate subject to result in prolonged suppression of reproductive
 CC behaviour and/or fertility. GnRH immunogens, analogues or antibodies are
 CC used to manufacture a composition or vaccine for immunosterilisation or
 CC immunoc contraception of feline, canine, equine or cervine subjects.
 CC The vaccines are used to suppress reproductive behaviour and/or
 CC fertility for at least 10 months. The prepubertal administration
 CC results in a prolonged, long-term suppression of testicular development
 CC and/or function in males, or a prolonged, long-term suppression of
 CC ovarian development and/or function in females. The methods provide a
 CC viable and desirable alternative to surgical forms of sterilisation that
 CC are currently used.

XX Sequence 49 AA:

alignment_scores:
 Quality: 290.00 Length: 49
 Ratio: 5.918 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-306-689-10 x Y58363 ..

Align seg 1/1 to: Y58363 from: 1 to: 49

1 CAGCATTTGAGCTACGCGCTCGCCCTGGCAGCGGTCTTCAGATTGGAG 50
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1 GlnH1StrpSerTyrGlyLeuArgProGlySerGlnAspTrpSe 17
 51 CTACGGCCTGCGTCCGGGTGCTCTAGCCAGCATTTGAGCTCGCCCTGC 100
 ||||||||||||||||||||||||||||||||||||||||||||||||
 17 rTyrGlyLeuArgProGlyGlySerSerGlnH1StrpSerTyrGlyLeuA 34
 101 GCCCTGGCAGCGGTAGCCAGATTGAGCTACGCGCTCGCCGGGT 147
 ||||||||||||||||||||||||||||||||||||||||||||||||
 34 rGProGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49
 seq_name: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:Y58135
 seq_documentation_block:
 ID Y58135 standard; Protein: 49 AA.
 XX
 AC Y58135;
 XX
 DT 07-MAR-2000 (first entry)
 XX
 DE GnRH analogue multimer, containing four copies of the GnRH analogue.
 XX
 KW Gonadotropin releasing hormone; GnRH; leukotoxin; LRT; fusion protein;
 KM antibody; immunogenic; chimeric; vaccine; testosterone; androgenic;
 KW non-androgenic; steroid; reduction; weight gain; muscle distribution;
 KW fat distribution; male pattern; boar taint; flavour; impairment;
 KW reliable; immunocastration; meat production.
 XX
 OS Synthetic.

```

OS Mammalia.
XX WO9956771-A2.
XX 11-NOV-1999.
XX
XX 05-MAY-1999; 99WO-CA00360.
XX
XX 05-MAY-1998; 98US-0084217.
XX
XX (BIOS-) BIOSTAR INC.
XX
XX Manns JG, Acres SD, Harland R;
XX
XX WPI; 2000-062125/05.
XX
XX N-PSDB; 246402.
XX
XX Production of uncastrated male food animals using vaccines -
XX
XX Example 1; Fig 2B; 87pp; English.
XX
XX This sequence represents four copies of a gonadotropin
XX releasing hormone (GnRH) analogue, DNA encoding which was
XX used in the construction of a chimeric GnRH-leukotoxin (LKT)
XX fusion gene (246400). This fusion gene encodes a GnRH-LKT fusion
XX protein which may be used as a vaccine. The LKT portion of the protein
XX acts to enhance the immunogenicity of the GnRH portion. The invention
XX relates to a method of using two GnRH immunogen vaccines to produce
XX uncastrated male animals for meat production, one vaccination prior to
XX or during the fattening period to reduce circulating testosterone levels,
XX and the second vaccination about 2-8 weeks before slaughter to
XX substantially reduce androgenic and/or non-androgenic steroids. The
XX invention is used to produce food animals that exhibit the weight gain
XX and muscle/fat distribution of male animals without the problems
XX associated with male animals. Such problems include "boar taint", a
XX urine-like odour found in cooked meat of uncastrated pigs which is
XX caused by steroids stored in the tissues, and similar flavour
XX impairments in the meat of other intact male animals. The invention is
XX more reliable than prior art immunocastration techniques.
XX
XX Sequence 49 AA:
SQ

```

alignment_scores:

Quality:	290.00	Length:	49
Ratio:	5.918	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

```

US-09-306-689-10 x Y58135
Align seg 1/1 to: Y58135 from: 1 to: 49

```

```

1 CAGCATTTGGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlySerGlnAspTrpSe 17
51 CTACGGCCTCGTCCGGGTGCTCTACGCAGATTGGAGCTAGCGCCTGC 100
17 rTyrglyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 34
101 GCCCTGGCAGCGGTAGCCAAAGATTGGAGCTACGGCTGCGCGGGT 147
34 rPrGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 49
seq_name: /STD1/gcgdata/geneseq/geneseqp/AA1996.DAT:W03943.
seq_documentation_block:
ID W03943 standard; Protein: 544 AA.
XX
XX W03943:
XX
XX 20-NOV-1996 (first entry)

```

```

XX DE LKT-GnRH protein fusion from PCB111.
XX
XX Leukotoxin; LKT: gonadotropin-releasing hormone; GnRH;
XX fusion protein; immunogen; vaccine; fertility control;
XX contraceptive; sterilisation.
XX
XX Chimeric Pasteurella haemolytica A1 strain B122;
XX Chimeric synthetic.
XX
XX Key Location/Qualifiers
XX Domain 1..493
XX FT /Label= LKT
XX FT 494..544
XX FT /Label= GnRH_repeat_domain
XX
XX WO9624675-A1.
XX
XX 15-AUG-1996.
XX
XX 24-JAN-1996; 96WO-CA00049.
XX
XX 10-FEB-1995; 95US-0387156.
XX
XX (UYSA-) UNITV SASKATCHEWAN.
XX
XX Manns JG, Potter AA;
XX
XX WPI; 1996-384447/38.
XX
XX N-PSDB; T37177.
XX
XX Gonadotropin-releasing hormone multimer fusion proteins - with
XX leukotoxin polypeptide for increased immunogenicity, useful in
XX antifertility vaccine prodn.
XX
XX Claim 8: Fig 7A-7E; 87pp; English.
XX
XX A chimeric protein (W03943) is composed of a fusion between
XX a truncated leukotoxin (LKT-111) from Pasteurella haemolytica
XX and a 4-copy gonadoliberin-releasing hormone (GnRH) repeat
XX sequence (see also W03944). It is the product of a chimeric
XX gene (T37177) produced by deleting an approx. 1300 bp sequence
XX from PCB113 (see also T37176) coding for amino acids 352-784
XX of LKT-352. Recombinant plasmid PCB111 (LKT 111:4 copy GnRH,
XX ATCC 69748) was obtd. Escherichia coli transformants produced
XX the chimeric protein, which is useful as a vaccine for fertility
XX control, esp. immunological sterilisation of domestic or farm
XX animals.
XX
XX Sequence 544 AA;
SQ

```

alignment_scores:

Quality:	290.00	Length:	49
Ratio:	5.918	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

```

US-09-306-689-10 x W03943
Align seg 1/1 to: W03943 from: 1 to: 544

```

```

1 CAGCATTTGGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
1 GlnHisTrpSerTyrGlyLeuArgProGlySerGlySerGlnAspTrpSe 510
494 GlnHisTrpSerTyrGlyLeuArgProGlySerGlySerGlnAspTrpSe 510
51 CTACGGCCTCGTCCGGGTGCTCTACGCAGATTGGAGCTAGCGCCTGC 100
510 rTyrglyLeuArgProGlySerGlnHisTrpSerTyrGlyLeuA 527
101 GCCCTGGCAGCGGTAGCCAAAGATTGGAGCTACGGCTGCGCGGGT 147
527 rPrGlySerGlySerGlnAspTrpSerTyrGlyLeuArgProGly 542

```

seq_name: /SIDS1/gcdata/geneseq/geneseqp/AA1998.DAT:W79570

seq_documentation_block:

ID W79570 standard; Protein; 544 AA.

XX W79570;

XX 24-DEC-1998 (first entry)

XX LKT-GnRH chimeric protein.

XX Chimera; PCB11; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;

XX cytotoxic activity; antigen presentation; immune response; vaccine;

XX tumour.

XX Synthetic.

XX WO9806848-A1.

XX 19-FEB-1998.

XX 08-AUG-1997; 97WO-CA00559.

XX 09-AUG-1996; 96US-0694865.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Manns JG, Potter AA;

XX WPI: 1998-159540/14.

XX N-PSDB; V61532.

XX Chimeric protein of leukotoxin and gonadotropin releasing hormone

XX useful for, e.g. preparation of vaccines for reduction of incidence

XX of mammary tumours in mammals

XX PS Disclosure; Figure 7.1-5; 118pp; English.

XX The present sequence represents the LKT-GnRH chimeric protein from

XX PCB11. This plasmid contains the LKT 111 polypeptide fused to

XX four copies of the GnRH peptide. This chimera lacks cytotoxic activity

XX which enables there to be an increase in antigen presentation and thus an

XX optimal immune response. The removal of this region also enables the

XX truncated LKT to be expressed at much higher levels and allows the amount

XX of antigen administered to be reduced. This chimeric protein comprises a

XX leukotoxin polypeptide, several multimers, and a GnRH sequence. The

XX chimeric protein can be used as a vaccine to help reduce the incidence of

XX mammary tumours in a mammalian individual.

XX SQ Sequence 544 AA;

alignment_scores:

Quality: 290.00

Ratio: 5.918

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-09-306-689-10 x W79570

Align seg 1/1 to: W79570 from: 1 to: 544

1 CAGCATGTGAGCTACGGCTCGCCCTGGACGGCTTCTCAAGATTGAG 50
|||||
494 GlnHisTSPSerTyrGlyLeuArpProGlySerGlySerGlnAspTyrPse 510
51 CTACGGCCCTGCGGCTCGGCTGCTAGCCAGCATTTGAGACTAGCGCTGC 100
|||||
510 rTyrglyLeuArpProGlyGlySerSerGlnHisTSPSerTyrGlyLeuA 527
|||||
101 GCCCTGACGACGGTAGCCAAAGATTGGAGCTACGGCTCGCGGGT 147
|||||

527 rProGlySerGlySerGlnAspTyrPseTyrGlyLeuArpProGly 542

seq_name: /SIDS1/gcdata/geneseq/geneseqp/AA1998.DAT:W79573

seq_documentation_block:

ID W79573 standard; Protein; 695 AA.

XX W79573;

XX 24-DEC-1998 (first entry)

XX LKT-GnRH chimeric protein.

XX Chimera; PCB122; LKT 111; GnRH; Gonadotropin releasing hormone; multimer;

XX cytotoxic activity; antigen presentation; immune response; vaccine;

XX tumour.

XX Synthetic.

XX WO9806848-A1.

XX 19-FEB-1998.

XX 08-AUG-1997; 97WO-CA00559.

XX 09-AUG-1996; 96US-0694865.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Manns JG, Potter AA;

XX WPI: 1998-159540/14.

XX N-PSDB; V61535

XX Chimeric protein of leukotoxin and gonadotropin releasing hormone

XX useful for, e.g. preparation of vaccines for reduction of incidence

XX of mammary tumours in mammals

XX PS Claim 9; Figure 9.1-6; 118pp; English.

XX The present sequence represents the LKT-GnRH chimeric protein from

XX PCB122. This plasmid contains the LKT 111 polypeptide fused to sixteen

XX copies of the GnRH peptide. In the pattern of: 8 copies of GnRH-LKT 111-8

XX copies of GnRH. This chimera lacks cytotoxic activity which enables

XX there to be an increase in antigen presentation and thus an optimal

XX immune response. The removal of this region also enables the truncated

XX LKT to be expressed at much higher levels and allows the amount of

XX antigen administered to be reduced. This chimeric protein comprises a

XX leukotoxin polypeptide, several multimers, and a GnRH sequence. The

XX chimeric protein can be used as a vaccine to help reduce the incidence of

XX mammary tumours in a mammalian individual.

XX SQ Sequence 695 AA;

alignment_scores:

Quality: 290.00

Ratio: 5.918

Percent Similarity: 100.000

Percent Identity: 100.000

alignment_block:

US-09-306-689-10 x W79573

Align seg 1/1 to: W79573 from: 1 to: 695

1 CAGCATGTGAGCTACGGCTCGCCCTGGACGGCTTCTCAAGATTGAG 50
|||||
9 GlnHisTSPSerTyrGlyLeuArpProGlySerGlySerGlnAspTyrPse 25
51 CTACGGCCCTGCGGCTCGGCTGCTAGCCAGCATTTGAGACTAGCGCTGC 100
|||||
25 rTyrglyLeuArpProGlyGlySerSerGlnHisTSPSerTyrGlyLeuA 42
|||||

alignment_scores:

Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x Y58133 ..

Align seg 1/1 to: Y58133 from: 1 to: 695

```

1 CACGATTGAGCTACGCGCTGCGCCCTGGCAGCGGTTCTCAAGATTGAG 50
|||||
9 GlnH1StrpSerTyrglyLeuArgProglySerGlySerGlnAspTrpse 25
25 rTyrglyLeuArgProglySerGlnAspTrpSerTyrglyLeuArgProgly 42
51 CTACGGCCTGCGTCCGGGTGCTTAGCCAGCATTTGAGAGCTACGCGCTGC 100
|||||
101 GCCCTGGCAGCGGTACGCCAAGATTGAGAGCTACGCGCTCGCGGT 147
|||||
42 rTyrglyLeuArgProglySerGlnAspTrpSerTyrglyLeuArgProgly 57
seq_name: /SIDSI/gcgsdata/geneseq/geneseq/AA1996.DAT:W03942

```

seq_documentation_block:

ID W03942 standard; Protein: 977 AA.

AC W03942:

DT 20-NOV-1996 (first entry)

DE LKT-GnRH protein fusion from PCB113.

XX Leukotoxin: LKT; gonadotropin-releasing hormone: GnRH;

KM fusion protein; immunogen; vaccine; fertility control;

KW contraceptive; sterilisation.

XX OS Chimeric Pasteurella haemolytica A1 strain B122;

OS Chimeric synthetic.

XX Key Location/Qualifiers

FF Domain 1..929

FT Domain /label= LKT

FT Domain 927..977

FT Domain /label= GnRH_repeat_domain

XX W09624675-A1.

PD 15-AUG-1996.

XX 24-JAN-1996; 96WO-CA00049.

XX 10-FEB-1995; 95US-0387156.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Manns JG, Potter AA;

DR WPI: 1996-384447/38.

DR N-PSDB: T37176.

PT Gonadotropin-releasing hormone multimer fusion proteins - with
leukotoxin polypeptide for increased immunogenicity, useful in
antifertility vaccine prodn.

XX Claim 7; Fig 5A-5H; 87pp; English.

CC A chimeric protein (W03942) is composed of a fusion between
a truncated leukotoxin (LKT-352) from Pasteurella haemolytica (see
also W03945) and a 4-copy gonadoliberin-releasing hormone (GnRH)
repeat sequence (see also W03944). It is the product of a

CC chimeric gene (T37176) produced by ligating a synthetic sequence
CC for the 4-copy GnRH into vector pAA352 (ATCC 68283), which carries
CC the LKT-352 gene. Recombinant plasmid PCB113 (LKT 352:4 copy
CC GnRH, ATCC 69749) was obt. Escherichia coli transformants
CC produced the chimeric protein, which is useful as a vaccine for
CC fertility control, esp. immunological sterilisation of
CC domestic or farm animals.

SQ Sequence 977 AA;

alignment_scores:

Quality: 290.00 Length: 49
Ratio: 5.918 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-306-689-10 x W03942 ..

Align seg 1/1 to: W03942 from: 1 to: 977

```

1 CACGATTGAGCTACGCGCTGCGCCCTGGCAGCGGTTCTCAAGATTGAG 50
|||||
927 GlnH1StrpSerTyrglyLeuArgProglySerGlySerGlnAspTrpse 943
51 CTACGGCCTGCGTCCGGGTGCTTAGCCAGCATTTGAGAGCTACGCGCTGC 100
|||||
101 GCCCTGGCAGCGGTACGCCAAGATTGAGAGCTACGCGCTCGCGGT 147
|||||
943 rTyrglyLeuArgProglySerGlnAspTrpSerTyrglyLeuArgProgly 960
960 rTyrglyLeuArgProglySerGlnAspTrpSerTyrglyLeuArgProgly 975
seq_name: /SIDSI/gcgsdata/geneseq/geneseq/AA1998.DAT:W79569

```

seq_documentation_block:

ID W79569 standard; Protein: 977 AA.

AC W79569;

DT 24-DEC-1998 (first entry)

DE LKT-GnRH chimeric protein.

XX Chimeric; PCB113; LKT 352; GnRH; Gonadotropin releasing hormone; multimer;

KM cytotoxic activity; antigen presentation; immune response; vaccine;

KW tumour.

XX OS Synthetic.

XX W09806848-A1.

PD 19-FEB-1998.

XX 08-AUG-1997; 97WO-CA00559.

XX 09-AUG-1996; 96US-0694865.

XX (UYSA-) UNIV SASKATCHEWAN.

XX Manns JG, Potter AA;

DR WPI: 1998-159540/14.

DR N-PSDB: V61531.

PT Chimeric protein of leukotoxin and gonadotropin releasing hormone
PT useful for, e.g. preparation of vaccines for reduction of incidence
PT of mammary tumours in mammals

XX Disclosure; Figure 5.1-8; 118pp; English.

CC The present sequence represents the LKT-GnRH chimeric protein from
CC PCB113. This plasmid contains the LKT 352 polypeptide (W79568) fused to

CC four copies of the GnRH peptide. This chimera lacks cytotoxic activity
 CC which enables there to be an increase in antigen presentation and thus an
 CC optimal immune response. The removal of this region also enables the
 CC truncated LRT to be expressed at much higher levels and allows the amount
 CC of antigen administered to be reduced. This chimeric protein comprises a
 CC leukotoxin polypeptide, several multimers, and a GnRH sequence. The
 CC chimeric protein can be used as a vaccine to help reduce the incidence of
 CC mammary tumours in a mammalian individual.

XX Sequence 977 AA:

alignment_scores:

Quality: 290.00 Length: 49
 Ratio: 5.918 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-306-689-10 x W79569 ..

Align seg 1/1 to: W79569 from: 1 to: 977

1 CAGCATTTGGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
 ||||||||||||||||||||||||||||||||||||||||||||||||
 927 GlnHstRPSerTYrGlyLeuAtrProGlySerGlySerGlnAspTrpSe 943
 51 CTACGGCGCTGCGGCTGCTAGCCAGCATTTGGAGCTACGGCGCTGC 100
 ||||||||||||||||||||||||||||||||||||||||||||||||
 943 rTYrGlyLeuAtrProGlySerGlySerGlnHstRPSerTYrGlyLeuA 960
 101 GCCCTGGCAGCGGTAGCCAGATTGGAGCTACGGCGCTGCGGCT 147
 ||||||||||||||||||||||||||||||||||||||||||||||||
 960 rGProGlySerGlySerGlnAspTrpSerTYrGlyLeuAtrProGly 975

seq_name: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.Y31183

seq_documentation_block:

ID Y31183 standard; peptide: 40 AA.

XX Y31183;

DT 28-OCT-1999 (first entry)

DE Ubiquitin fusion protein GnRH fragment 2.

XX Ubiquitin: immunocastration; fusion protein; heat shock protein; epitope;
 KW immune response stimulation; vaccine; T cell; viral; infection; cancer;
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;
 KW pulmonary infection; respiratory infection; scaffold; anti-self; pig;
 KW steroidogenesis; gamete maturation; prostate; breast; castration; TNF;
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;
 KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;
 KW fertility; sperm protein; growth rate; antibody; detection; GnRH.

XX unidentified.

XX OS

XX MO942472-AL.

XX 26-AUG-1999.

XX 26-JAN-1999; 99MO-US01588.

XX 19-FEB-1998; 98US-0026276.

XX (IGEN-) IGEN INT INC.

XX Kenten JH, Lohmas GL, Pilon AL, Roberts SF, Tramontano A;

XX WPI; 1999-518582/43.

PT Epitope-containing fusion proteins used to generate a highly
 PT specific immune responses
 XX

PS Claim 83; Page 43; 67p; English.

XX This invention describes a novel fusion protein, comprising a heat shock
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner
 CC which is useful for the stimulation of a highly specific immune response
 CC when administered to an animal. The protein of the invention may be
 CC post-translationally modified (e.g. by the addition of fatty acids to
 CC enhance immunogenicity). The fusion proteins of the invention can be
 CC used as vaccines to induce an immune response. When a T cell epitope is
 CC attached, they can be used for control of viral infections, bacterial
 CC infections, parasitic infection and cancer. The fusion proteins can be
 CC used in pharmaceutical compositions for the treatment of gastrointestinal
 CC diseases, pulmonary infections, respiratory infections, and HIV
 CC infections. The use of ubiquitin as a scaffold is also useful for the
 CC presentation and stimulation of anti-self immune responses, e.g.
 CC generation of anti-gonadotropin releasing hormone antibodies which result
 CC in the suppression of luteinizing hormone and follicle stimulating
 CC hormone. This indirectly suppresses steroidogenesis and gamete maturation
 CC in males and females. This type of anti-self response in humans is useful
 CC in the treatment of prostate cancer and breast cancer. In livestock, the
 CC ability to stimulate an anti-self response provides a simple alternative
 CC to physical castration. Immunocastration of pigs is a better alternative
 CC to physical castration, as it does not result in any of the detrimental
 CC side effects associated with physical castration. Other examples of
 CC diseases and conditions treated with self proteins fused with ubiquitin
 CC are TNF and its epitopes to modulate septic shock, arthritis,
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig
 CC gonadotropin for fertility control; and sperm proteins for fertility
 CC control. A further use of the fusion proteins is as part of a vaccine to
 CC enhance growth rate and thereby the final weight of the livestock prior
 CC to shipment to market. In addition, the fusion proteins of the invention
 CC can be used to detect and identify antibodies from experimental samples.
 CC This sequence represents a GnRH fragment used in the construction of
 CC a ubiquitin fusion protein described in the method of the invention.

XX Sequence 40 AA:

alignment_scores:

Quality: 199.50 Length: 49
 Ratio: 5.250 Gaps: 3
 Percent Similarity: 77.551 Percent Identity: 77.551

alignment_block:

US-09-306-689-10 x Y31183 ..

Align seg 1/1 to: Y31183 from: 1 to: 40

1 CAGCATTTGGAGCTACGGCTGCGCCCTGGCAGCGGTTCTCAAGATTGGAG 50
 ||||||||||||||||||||||||||||||||||||||||||||||||
 1 GlnHstRPSerTYrGlyLeuAtrProGly.....GlnHstRPS 14
 51 CTACGGCGCTGCGGCTGCTAGCCAGCATTTGGAGCTACGGCGCTGC 100
 ||||||||||||||||||||||||||||||||||||||||||||||||
 14 rTYrGlyLeuAtrProGly.....GlnHstRPSerTYrGlyLeuA 28
 101 GCCCTGGCAGCGGTAGCCAGATTGGAGCTACGGCGCTGCGGCT 147
 ||||||||||||||||||||||||||||||||||||||||||||||||
 28 rGProGly.....GlnHstRPSerTYrGlyLeuAtrProGly 40

seq_name: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.Y31182

seq_documentation_block:

ID Y31182 standard; peptide: 41 AA.

XX Y31182;

DT 28-OCT-1999 (first entry)

DE Ubiquitin fusion protein GnRH fragment.

XX Ubiquitin: immunocastration; fusion protein; heat shock protein; epitope;
 KW

KW Immune response stimulation; vaccine; T cell; viral; infection; cancer;
 KW bacterial; parasitic; treatment; gastrointestinal disease; HIV infection;
 KW pulmonary infection; respiratory infection; scaffold; anti-self; pig;
 KW steroidogenesis; gamete maturation; prostate; breast; castration; TNF;
 KW tumour necrosis factor; septic shock; arthritis; Crohn's disease;
 KW inflammatory bowel disease; ulcerative colitis; chorionic gonadotropin;
 KW fertility; sperm protein; growth rate; antibody; detection; GnRH.
 XX
 OS Unidentified.
 XX
 PM MO9942472-A1.
 XX
 PD 26-AUG-1999.
 XX
 PF 26-JAN-1999; 99WO-US01588.
 XX
 PR 19-FEB-1998; 98US-0026276.
 XX
 PA (IGEN-) IGEN INT INC.
 XX
 PI Keuten JH, Lohmas GL, Pilon AL, Roberts SF, Tramontano A;
 DR WPI: 1999-518582/43.
 XX
 PT Epitope-containing fusion proteins used to generate a highly
 PS specific immune responses
 XX
 PS Claim 01; Page 43; 67pp; English.
 CC This invention describes a novel fusion protein, comprising a heat shock
 CC protein (e.g. ubiquitin), fused to an epitope(s) in a defined manner
 CC which is useful for the stimulation of a highly specific immune response
 CC when administered to an animal. The protein of the invention may be
 CC post-translationally modified (e.g. by the addition of fatty acids to
 CC enhance immunogenicity). The fusion proteins of the invention can be
 CC used as vaccines to induce an immune response. When a T cell epitope is
 CC attached, they can be used for control of viral infections, bacterial
 CC infections, parasitic infection and cancer. The fusion proteins can be
 CC used in pharmaceutical compositions for the treatment of gastrointestinal
 CC diseases, pulmonary infections, respiratory infections, and HIV
 CC infections. The use of ubiquitin as a scaffold is also useful for the
 CC presentation and stimulation of anti-self immune responses, e.g.
 CC generation of anti-gonadotropin releasing hormone antibodies which result
 CC in the suppression of luteinizing hormone and follicle stimulating
 CC hormone. This indirectly suppresses steroidogenesis and gamete maturation
 CC in males and females. This type of anti-self response in humans is useful
 CC in the treatment of prostate cancer and breast cancer. In livestock, the
 CC ability to stimulate an anti-self response provides a simple alternative
 CC to physical castration. Immunocastration of pigs is a better alternative
 CC to physical castration, as it does not result in any of the detrimental
 CC side effects associated with physical castration. Other examples of
 CC diseases and conditions treated with self proteins fused with ubiquitin
 CC are TME and its epitopes to modulate septic shock, arthritis,
 CC inflammatory bowel disease, Crohn's disease, and ulcerative colitis; Ig
 CC epsilon heavy chain for the control of allergic reactions; chorionic
 CC gonadotropin for fertility control; and sperm proteins for fertility
 CC control. A further use of the fusion protein is as part of a vaccine to
 CC enhance growth rate and thereby the final weight of the livestock prior
 CC to shipment to market. In addition, the fusion proteins of the invention
 CC can be used to detect and identify antibodies from experimental samples.
 CC This sequence represents a GnRH fragment used in the construction of
 CC a ubiquitin fusion protein described in the method of the invention.
 XX
 XX
 50 Sequence 41 AA:

alignment_scores:

Quality: 199.50

Ratio: 5.250

Percent Similarity: 77.551

Length: 49

Gaps: 3

Percent Identity: 77.551

alignment_block:

US-09-306-689-10 x Y31182

```

Align seg 1/1   to: Y31182    from: 1   to: 41

      1 CAGCAATTGGACCTACGCGCCTGCSCGCCGTTCGAAGAATTTGAG 50
        |||||||
      1 GlnHisTrpSerTyrGlyLeuArgProGly.....GlnHisTrpSer 14
      14 TTYcGlYleuArPrOgLy.....GlnHisTrpSerTyrGlyLeuA 28
      101 GCCCTGGCGCCGCTAGCCAAAGATTGGAGCTACGCGCCTGCCTGGCGGT 147
          |||||
      28 rgrProGly.....GlnHisTrpSerTyrGlyLeuArProGly 40

seq_name: /STDS1/gcgdata/geneseq/geneseqp/AAL191.DAT:R1187
seq_documentation_block:
ID R1187 standard; Protein; 323 AA.
XX
AC R1187;
XX
DT 22-MAY-1991 (first entry)
DE Plasmid pBTA859-encoded TrtRp-multiple LHRN analogue fusion.
XX
KW TrtRp protein; Leutinizng hormone releasing hormone; fusion protein;
  Immunological castration.
XX
FH Key Location/Qualifiers
FT Peptide 1..20
FT Peptide /label= TrtRp signal
FT Peptide 201..280
FT Peptide /label= 8 LHRN analogues in tandem repeat
PN WO9102799-A.
PP PD 07-MAR-1991.
PF PF 24-AUG-1990; 90WO-AU00373.
XX XX 25-AUG-1989; 89AU-0005979.
PA (BIOR-) BIOTECHN AUST Pty L.
PI Russell-Jones GJ, Stewart AG, Tsomis CG;
XX NPSI; 1991-087282/12.
DR DR N-PDB; Q11021.
XX
PT Fusion proteins comprising LHRN analogue and TrtRp (analogue) -
PT useful in vaccine for inhibition or control of reproduction in
PT vertebrates, esp. domestic animals
XX
XX Example 1; Fig 2 and 5; 53pp; English.
PS
XX Plasmid pBTA859 is a TrtRp-LHRN analogue fusion in which 8 copies
  of an LHRN analogue have been inserted between amino acids 200 and
  CC 201 of TrtRp (Oyata R.T. et al., (1982) J.Bacteriol. 151:819-827).
  CC The plasmid was constructed by two successive additions of DNA
  CC coding for a dimer of LHRN analogue into the SmaI site of pBTA862
  CC (see Q11020) which all ready carries four copies of the LHRN
  CC sequence. After transformation, colonies with 8 LHRN molecules were
  identified. Fusion proteins with multiple inserts generated a higher
  anti-LHRN response (as measured by the binding of (125)I-LHRN at a
  serum dilution of 1:2000 final) than constructs with a single
  insert, in outbred mice and dogs. The fusion proteins can be used to
  inhibit reproductive functions in vertebrates.
  See also Q10995, Q10997-Q11000, Q11014-Q11020.
XX
Sequence 323 AA:
```

alignment_scores:

Quality: 187.50 Length: 49
 Ratio: 4.934 Gaps: 3
 Percent Similarity: 77.551 Percent Identity: 69.388

alignment_block:

US-09-306-689-10 x R11187 ..

Align seg 1/1 to: R11187 from: 1 to: 323

```

1 CAGCATTGGAGCTACGGGCTGCGCCCTGGCAGCGGTTCTCAGATTGGAG 50
  :::::::::::::::::::::::::::::::::::::
201 GluH1StrpSerTyrglyLeuArgProgly.....GluH1StrpSe 214
51 CTACGGGCTGCGGCTGCGCTAGCCAGCATTGGAGCTAGCGGCTGC 100
  :::::::::::::::::::::::::::::::::::::
214 rTyrglyLeuArgProgly.....GluH1StrpSerTyrglyLeuA 228
101 GCCCTGGCAGCGGTAGCCAGAGATTGAGCTAGGCTGCGCTCCGGGT 147
  :::::::::::::::::::::::::::::::::::::
228 rgProgly.....GluH1StrpSerTyrglyLeuArgProgly 240
    
```